Machine Learning and Pattern Recognition Methods for Remote Sensing Image Registration and Fusion

David Solarna

Advisor: Prof. Gabriele Moser

A thesis presented for the degree of
Doctor of Philosophy

University of Genoa

Department of Electric, Electronic, and Telecommunication Engineering and Naval Architecture

January, 30, 2021

Contents

List of Figures

List of Tables

Abstract

In the last decade, the remote sensing world has dramatically evolved. New types of sensor, each one collecting data with possibly different modalities, have been designed, developed, and deployed. Moreover, new missions have been planned and launched, aimed not only at collecting data of the Earth's surface, but also at acquiring planetary data in support of the study of the whole Solar system. Indeed, such a variety of technologies highlights the need for automatic methods able to effectively exploit all the available information. In the last years, lot of effort has been put in the design and development of advanced data fusion methods able to extract and make use of all the information available from as many complementary information sources as possible. Indeed, the goal of this thesis is to present novel machine learning and pattern recognition methodologies designed to support the exploitation of diverse sources of information, such as multisensor, multimodal, or multiresolution imagery. In this context, image registration plays a major role as is allows bringing two or more digital images into precise alignment for analysis and comparison. Here, image registration is tackled using both feature-based and area-based strategies. In the former case, the features of interest are extracted using a stochastic geometry model based on marked point processes, while, in the latter case, information theoretic functionals and the domain adaptation capabilities of generative adversarial networks are exploited. In addition, multisensor image registration is also applied in a large scale scenario by introducing a tiling-based strategy aimed at minimizing the computational burden, which is usually heavy in the multisensor case due to the need for information theoretic similarity measures. Moreover, automatic change detection with multiresolution and multimodality imagery is addressed via a novel Markovian framework based on a linear mixture model and on an ad-hoc multimodal energy function minimized using graph cuts or

belied propagation methods. The statistics of the data at the various spatial scales is modelled through appropriate generalized Gaussian distributions and by iteratively estimating a set of virtual images, at the finest resolution, representing the data that would have been collected in case all the sensors worked at that resolution. All such methodologies have been experimentally evaluated with respect to different datasets, and with particular focus on the trade-off between the achievable performances and the demands in terms of computational resources. Moreover, such methods are also compared with state-of-the-art solutions, and are analyzed in terms of future developments, giving insights to possible future lines of research in this field.

1

An Introduction to Remote Sensing

1.1 Remote Sensing Imagery

The field of remote sensing usually refers to the discipline aiming at retrieving information about a given "object" by exploiting observation data acquired by sensors which are not in physical contact with the object itself [1]. More specifically, the “object” taken into account by remote-sensing methodologies dealing with Earth-Observation (EO) applications is typically a given geographical area of interest. It includes, but is not limited to, the mapping of land cover, land use, and of geophysical or biophysical properties of the observed surface, statically at a single time or dynamically along a temporal series. Indeed, what is “of interest” (or what is not) is an application-dependent dilemma [2].

This type of technology has been acquiring a growing interest from the viewpoints of environmental monitoring and management, thanks to the repetitive geographical area coverage it provides to the end user. In fact, thanks to the increasing number of missions devoted to putting EO satellites in orbit and to the corresponding growing availability of extensive EO imagery, the remote-sensing technology presents huge potentialities for environmental applications both on a global scale (e.g., global warming monitoring or analysis of changes in the Earth system) and on a regional-scale (e.g.,

weather forecasting and modelling, desertification studies), and on a local scale (e.g., crop monitoring, urban-areas mapping). More generally, remote-sensing data can provide valuable information for a large variety of applications, from vegetation-resource management and ecology (e.g., precision farming, forest mapping and inventory) to urban and land-use applications, from meteorological, oceanographic, and hydrological applications (e.g., ice, snow-cover, and drought monitoring, water-quality assessment) to geological (e.g., stratigraphy studies) and geophysical (e.g., crustal-dynamic monitoring, Earth magnetic field studies) applications. In addition, remote sensing can also provide a crucial support for handling natural disasters (e.g., forest fires, floods, landslides, earthquakes, or seaquakes), both for prevention purposes (e.g., through the generation of risk maps for flooding events) and as a support to crisis management and to post-crisis damage assessment (e.g., through the generation of maps of burnt areas after a fire event).

The remote sensor is usually carried by airborne or spaceborne platforms, and the data set acquired by the sensor is typically expressed as a digital (either scalar-valued or vector-valued) image of the area of interest [1] (for additional details refer to Section 1.1.1). From this viewpoint, the goal of the remote-sensing technology turns out to be the extraction from such image of thematic, geophysical, and geographical information being of interest to the end-user (see Figure 1.1).

Indeed, the following sub-sections will provide additional details on the platforms deployed to carry the remote sensors, including the airborne and spaceborne categorization, the different types of orbit and their characteristics, and will also provide details on the acquired digital images and the underlying concepts that will be the basis for the algorithms and methods addressed in later Chapters.

1.1.1 Digital Images, Platforms and Orbits

Practically speaking, remote sensing images are 2D tables of points, named pixels (abbreviation of "picture elements"), which are associated with one or more discrete values (pixel intensities) [3]. Given a digital image collected by a certain sensor, the meaning of the pixel intensities in terms of measurements of physical quantities substantially depend on the sensor itself [4].

From a signal-processing viewpoint, an image can be viewed as a real-

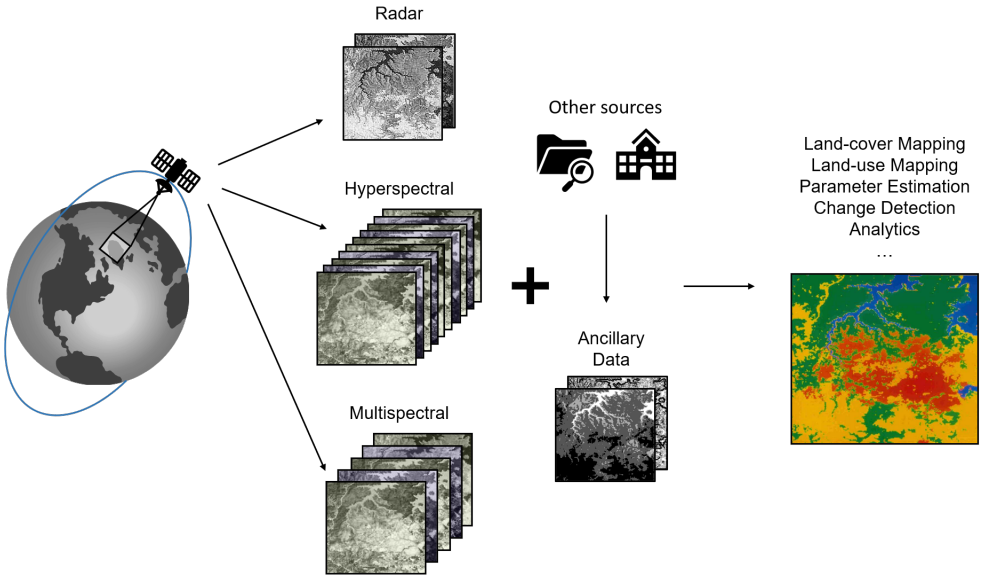


Figure 1.1: An overview of remote sensing data processing.

ization of a 2D stochastic process (also often named random field) defined on the discrete lattice of the pixel grid [5]. This perspective is especially convenient whenever probabilistic and statistical modeling are necessary, a frequent situation when processing and analysis tasks have to be addressed.

Other than in the pixel lattice, remote sensing data is also discrete in terms of measured values. Indeed, when collected by the sensor, remote sensing images always undergo a digitization process. Scalar-valued and vector-valued random fields are used in the cases of a unique intensity or of multiple intensities associated with each pixel, respectively. From a computational standpoint, a scalar-valued image is actually a rectangular matrix whose numbers of rows and columns correspond to the height and width of the image, respectively. Analogously, a vector-valued image can be pictured as a data cube, whose sizes correspond to the width, the height, and the number of components of the vector-valued pixel intensities (a third-order tensor) [6].

As anticipated in Section 1.1, the platforms carrying the actual imaging sensors can be broadly categorized as airborne and spaceborne. Major examples of airborne platforms include aircrafts, balloons, and unmanned aerial

vehicles (UAVs), also referred to as drones [1]. Aircrafts that can be equipped with EO sensors are usually available at national or international organizations (e.g., military authorities) and at specialized companies. Balloons are overall less frequently employed for EO, but their use is steadily growing and may probably increase in the future. UAVs and drones have been getting increasingly popular lately because of their low cost and, especially in case of drones, their capability to fly at lower altitude and in more cluttered scenarios. Nevertheless, as a drawback, they usually exhibit limitations on the maximum weight of the sensors they can carry. With both aircrafts and UAVs, altitude and orientation affect the geometry of the image, and acquisitions occur through ad-hoc flights. Helicopters are also sometimes used for airborne remote sensing, especially for specific applications.

A spaceborne platform is generally an artificial satellite orbiting around the Earth [1]. Exceptions include missions, such as the Spaceborne Imaging Radar-C / X-band Synthetic Aperture Radar (SIR-C/X-SAR) and the Shuttle Radar Topography Mission (SRTM), in which sensors were put on-board of the NASA Space Shuttles [7]. Using the language of satellite missions, in a spaceborne EO system, the satellite represents the space segment while the ground segment is the infrastructure on the Earth surface that receives, validates, and pre-processes the acquired data. The space segment of an EO mission includes the mission payload, i.e., the sensor(s) that the satellite is designed to carry, and all necessary infrastructures for power, orbit management, on-board pre-processing, recording, and transmission to the ground segment [8].

The path of a satellite along its orbit is an ellipse [1]. The plane that includes this ellipse is named orbital plane. The orbits used for EO are geostationary or near polar. A geostationary orbit (or geosynchronous equatorial orbit, GEO) is circular, its orbital plane is the plane of the Earth Equator, and the orbiting period around the Earth is 24 hours [9]. Therefore, a sensor on-board a geostationary platform (geostationary sensor) always observes the same portion of the Earth surface. Simple calculations based on Newton's gravitation law imply that the altitude of a geostationary orbit is approximately 36000 km above the Earth Equator (for comparison purposes, recall that the mean Earth radius is estimated as 6371 km). Weather satellites (e.g., Meteosat Second Generation, MSG) are most often geostationary.

In the case of a near polar orbit, the angle between the orbital plane

and the plane of the Equator is approximately 80° - 85° (or equivalently 95° - 100°). Note also that, by convention, angles greater than 90° refer to retrograde orbits, that are typical for sun-synchronous Earth observation satellites. Altitude is generally 400 - 1000 km, a range that is included within the broader family of low Earth orbits (LEO) [9]. Owing to the combination of the motion of the satellite around the Earth and of the rotation of the Earth itself on its axis, a sensor on-board a near polar satellite (near polar sensor) collects data over almost all the Earth surface. The projection of the satellite path on the Earth surface is usually named satellite ground track.

A special and very often used case of near polar orbit is the Sun-synchronous orbit (also known as heliosynchronous orbit). In this case, the angle between the orbital plane and the segment joining the centers of the Earth and the Sun is nearly constant in time [9]. Therefore, a sensor on-board a Sun-synchronous satellite (Sun-synchronous sensor) observes a given ground area at approximately the same time of the day on each consecutive overpass. This contributes to minimizing the differences in Sun illumination conditions across different observation times.

1.2 Active and Passive Remote Sensing

Remote sensors for EO applications can be broadly categorized into two classes, i.e., passive sensors and active sensors. In the case of active sensors, a signal is transmitted toward the considered surface and the resulting "echo" is backscattered and measured. Conversely, in the case of passive sensors, no signal is transmitted, but the radiation coming from the considered surface is directly received and measured (see Figure 1.2).

1.2.1 Passive Sensors

According to [2], a passive sensor receives the electromagnetic radiation that comes from the considered portion of the Earth surface either because it originates from the reflection of incident solar radiation or because it is spontaneously emitted by the surface itself.

The physical quantity measured by a passive EO sensor is the spectral radiance (or specific intensity). It is a radiometric quantity that characterizes the distribution of radiation in space and represents the power per unit of

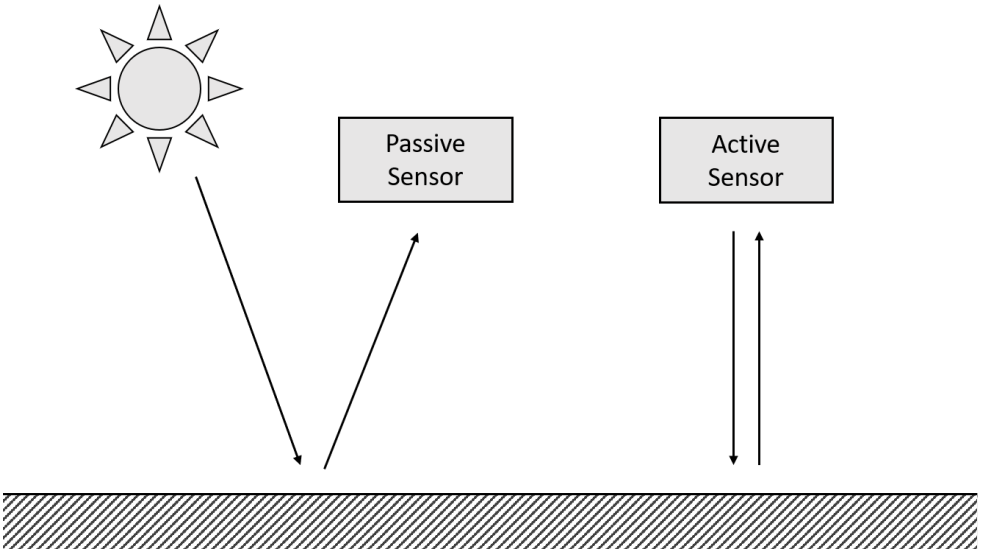


Figure 1.2: Difference between active and passive sensors [10].

wavelength that travels in a unitary solid angle centered on a given direction through a unitary surface, and it is measured in $[\mathrm{W} \cdot \mathrm{m}^{-2} \cdot \mathrm{sr}^{-1} \cdot \mu\mathrm{m}^{-1}]$ [9].

In the case of radiation in the visible portion of the electromagnetic spectrum (i.e., with wavelength between approximately 0.4 and 0.7 μm), in the near infrared range (NIR, 0.7 - 1.1 μm), and in the short-wave infrared range (SWIR, 1.1 - 1.35 μm, 1.4 - 1.8 μm, and 2 - 2.5 μm), spontaneous thermal emission from the Earth surface is negligible as compared to reflected solar radiation [11]. Therefore, the spectral radiance received by a passive sensor operating in these ranges depends on the reflective properties of the observed surface.

Vice versa, in the case of radiation in the thermal infrared (TIR, also known as long-wave infrared, LWIR) portion of the spectrum (i.e., approximately 8 - 9.5 μm and 10 - 14 μm), the reflected solar radiation is negligible as compared to the Earth thermal emission [11]. Therefore, the received radiance depends on the properties of the observed surface that characterize its capability to spontaneously emit radiation. Because of the well-known Planck's law, these quantities include the surface temperature $[K]$ and emittance [adimensional] [11]. In the intermediate case of mid-wave infrared radiation (MWIR, i.e., around 3 - 4 μm and 4.5 - 5 μm), reflection-based and

emission-based contributions are comparable, and their reciprocal weights generally depend on all the aforementioned surface properties.

In all such cases, the spectral radiance that reaches the sensor first has to propagate through the portion of the atmosphere that is in between the surface and the sensor itself. In the case of reflected solar radiation, propagation through the atmosphere occurs twice, first downward from the direction of the Sun along the path from the top of the atmosphere to the surface and then upward from the surface to the sensor. Propagation through the atmosphere, which is composed of a large number of particles, affects a spectral radiance field due to: (i) the thermal emission of radiation by the atmosphere itself; and (ii) the extinction of the propagated radiance field due to absorption (i.e., conversion of part of the energy associated with the radiation to heat) and to scattering from one propagation direction to another (i.e., redistribution of the energy associated with the radiation through different directions) [12]. These phenomena are quantitatively well described by the so-called radiative transfer equation, which is an integro-differential equation that can be explained in terms of conservation of energy and of scattering in random media [12]. The solution is generally a complex problem for which specific numerical techniques have been developed [12].

Once the physical quantities that passive remote sensing refers to has been introduced, it is worth briefly describing the imaging technologies that have been developed and that are commonly used for passive acquisitions. Those range from traditional cameras, to scanners that record images of the Earth's surface by moving the instantaneous field of view (IFOV) of the instrument across the surface to record the upwelling energy. The forward motion of the vehicle allows an image strip to be built up from the raster scans, and the portion of the ground (in the across-track direction) that is observed while the platform travels his orbit is called swath width and it is determined by the field of view (FOV) of the sensor [1] (Figure 1.3).

With the availability of reliable detector arrays based on charge coupled device (CCD) technology, an alternative and more recent image acquisition mechanism utilises what is commonly called a "push-broom" technique. In this approach a linear CCD imaging array is carried on the satellite normal to the platform motion as shown in Figure 1.4. As the satellite moves forward, the array records a strip of image data, equivalent in width to the field of view seen by the array. Additionally, each individual detector records a strip

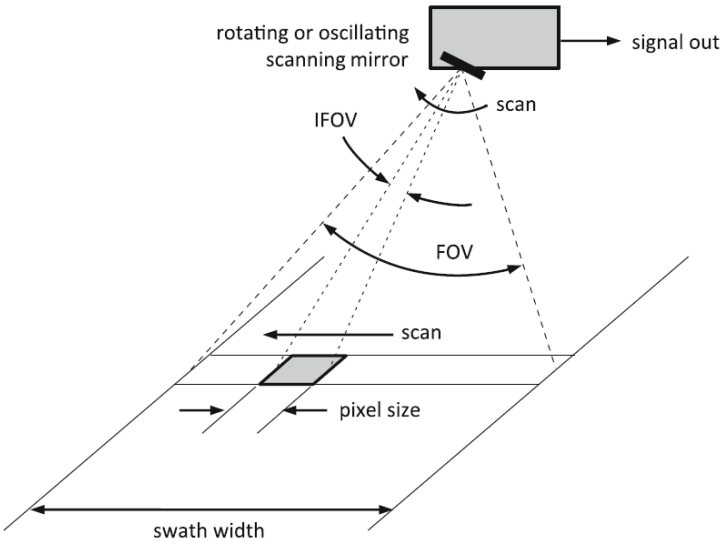


Figure 1.3: The image acquisition process by mechanical line scanning [1].

that is equivalent, in width, to the size of a pixel. Because of the larger time that is available (per pixel) for the integration of the energy emanating from the Earth's surface, the spatial resolution that can be achieved with push broom technology is better than with mechanical scanners [1].

Two dimensional CCD arrays are also available and find application in satellite imaging sensors. However, rather than recording a two-dimensional snapshot image of the earth's surface, the array is employed in a push broom manner. The second dimension is used to record simultaneously a number of different wavebands for each pixel. Such an arrangement is shown in Figure 1.5. Such devices are often referred to as imaging spectrometers, and the resulting data is referred to as multispectral or hyperspectral imaging, depending on number and the bandwidth of the channels that are recorded [1]. For additional details on multispectral and hyperspectral imagery refer to Section 1.3.

As anticipated in the previous paragraph, a passive EO sensor is most often designed to be multispectral, i.e., it collects data simultaneously from multiple wavelength ranges, named bands or channels. In particular, one speaks of a hyperspectral sensor if a large number (usually hundreds) of

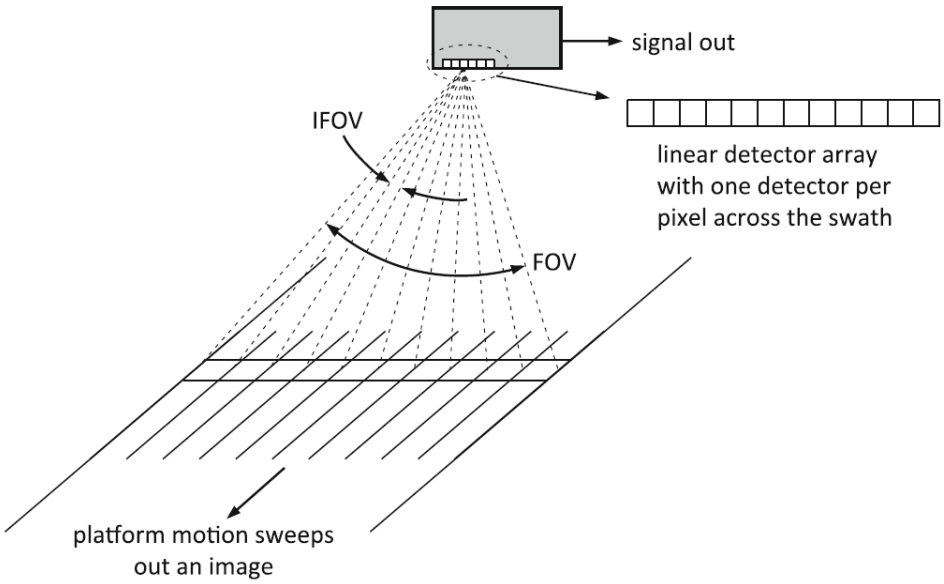


Figure 1.4: The image acquisition process by push-broom scanning [1].

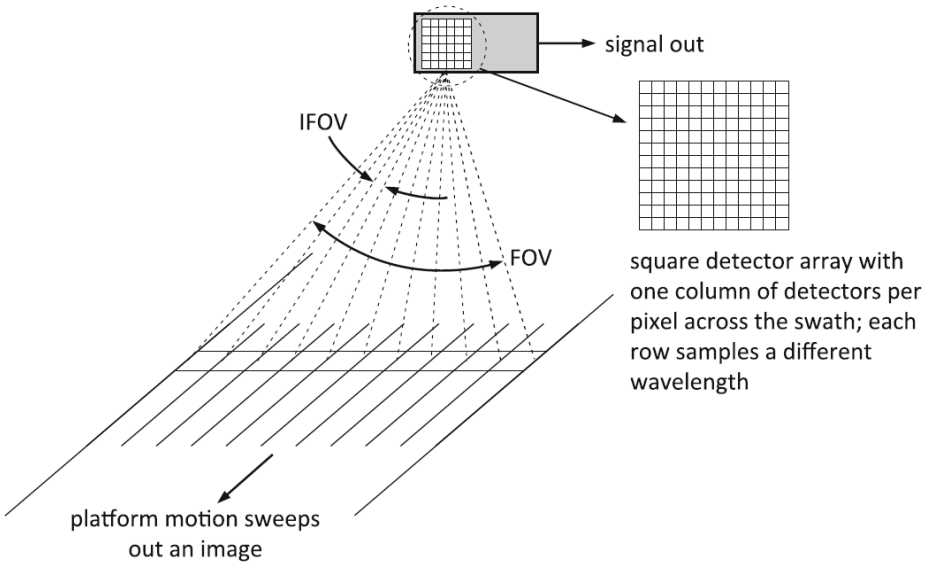


Figure 1.5: The image acquisition process by push-broom scanning with an array that allows the recording of several wavelengths simultaneously [1].

channels with narrow bandwidths are collected. Vice versa, a sensor designed to acquire only one channel, which usually encompasses the whole visible (and possibly NIR) range, is named panchromatic [13].

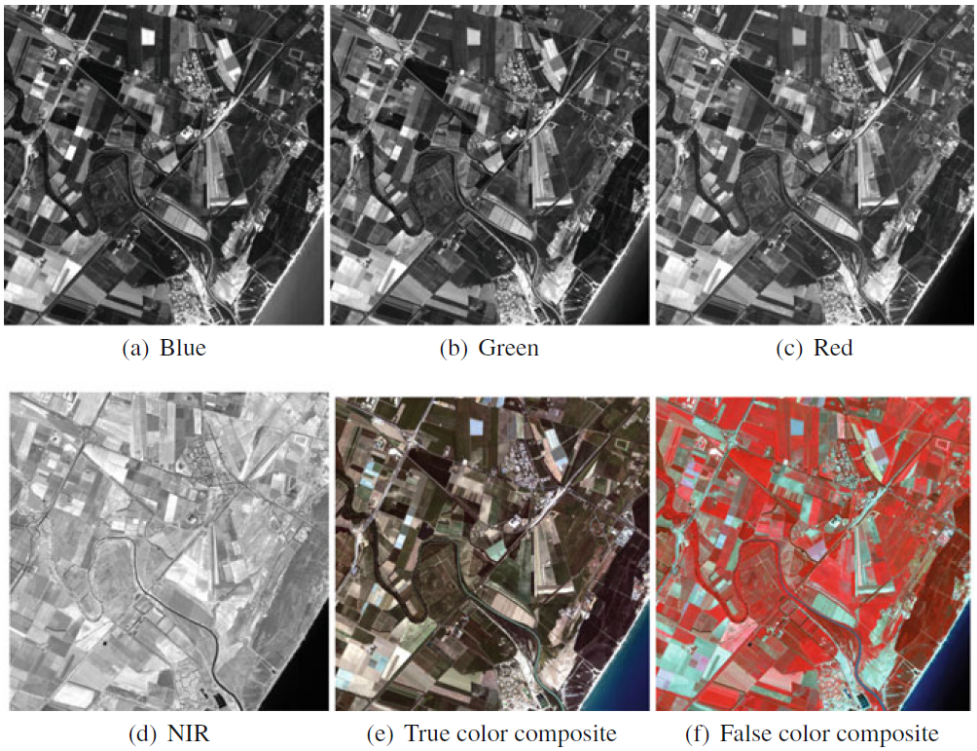
Multispectral acquisition can be accomplished using prisms and optical filters, which split the incoming radiance into different wavelength ranges, or using separate cameras that operate in distinct wavelength ranges directly. Figure 1.6 shows a multispectral image acquired in 2004 by the IKONOS sensor over Metaponto, Italy. The image is composed of four channels, approximately corresponding to the blue, green, red, and NIR wavelength ranges. Examples of color composites, in which the R, G, and B components of a displayed color image are associated with three of the available channels of the multispectral remote sensing images, are also shown in Figure 1.6(e) and (f). It is worth noting that, because of the aforementioned physical processes that lead to image formation, data collected by passive sensors are obviously affected by atmospheric (e.g., cloud cover) and Sun-illumination conditions.

1.2.2 Active Sensors

According to [2], an active sensor transmits an electromagnetic pulse toward the considered portion of the Earth surface and receives the resulting "echo" signal. For the purpose of 2D remote sensing image acquisition, microwave signals are typically used, and the imaging system is based on a radar (RAdio Detection And Ranging) instrument [1].

It is worth noting that a laser source can also be used for active remote sensing in a LiDAR (Light Detection And Ranging) instrument, also known as airborne laser scanning (ALS) or LaDAR (Laser Detection And Ranging) [14]. LiDAR is one of the most prominent technology for 3D mapping through remote sensing. However, since this type of data will not be tackled in this thesis, the LiDAR technology will not be discussed any further. For additional details refer to [14].

A radar imager for EO periodically emits a short-duration microwave pulse that is irradiated by a directive antenna as an electromagnetic wave in space. Part of the irradiated energy hits the considered surface that re-irradiates it in multiple directions, a phenomenon known as scattering [12]. The portion of the re-irradiated signal that is backscattered in the direction of the antenna is received by the antenna itself. In the application of radar to



(a) Blue (b) Green (c) Red

(d) NIR (e) True color composite (f) False color composite

Figure 1.6: Example of a multispectral image acquired by the passive IKONOS sensor over Metaponto, Italy (1250 × 1250 pixels). The different panels correspond to: (a) blue, (b) green, (c) red, and (d) NIR radiation; (e) the true color composite, in which the R, G, and B components of the displayed image are associated with the red, green, and blue channels of the multispectral image, respectively; and (f) a false color composite, in which the R, G, and B components of the displayed image are associated with the NIR, red, and green channels of the multispectral image, respectively.

positioning, the backscattered signal can be used for detecting the presence of a given target object (e.g., an aircraft), for measuring the distance of this target through the time taken by the pulse to reach the target and get back to the antenna and for estimating the speed of the target through the Doppler effect [15].

In the case of remote sensing image acquisition, the antenna is put onboard an airborne or spaceborne platform, and the goal is to use the backscattered signal to measure electromagnetic properties of the considered portion of the Earth surface in the microwave range. In the basic configuration of a single-frequency and single-polarization radar system for EO, the main property that is measured is the backscattering coefficient [adimensional], which is related to the average power of the return signal [15]. It is affected by numerous factors, including the roughness of the surface, its moisture content if it is a soil area, the presence on the surface of 3D structures (e.g., buildings), the carrier frequency of the microwave pulse, and the radar polarization.

Regarding the carrier frequency and the corresponding wavelength, we recall that, in general, the word "microwave" broadly refers to electromagnetic waves with frequency between 1 and 100 GHz, although precise definitions may vary. Specifically, using the IEEE Std 521 standard for radar frequency bands, we can mention the L-band (i.e., 1 - 2 GHz of carrier frequency or equivalently 15 - 30 cm of wavelength), the C-band (i.e., 4 - 8 GHz or 3.75 - 7.5 cm), and the X-band (i.e., 8 - 12 GHz or 2.5 - 3.75 cm) among the most common ranges for radar EO.

The pulse signal used by a radar for EO exhibits a narrowband spectrum in a neighborhood of the carrier frequency and is most usually a linearly frequency-modulated signal known as chirp [1]. This choice, together with appropriate filtering of the return signal, makes it possible to achieve high spatial resolution along the looking direction of the radar (named the range direction) [15].

A further radar acquisition technique is the so called synthetic aperture radar (SAR), which makes use of the motion of the platform along its path to simulate a long antenna, which, in turn, makes it possible to achieve high spatial resolution along the flight direction (named the azimuth direction) from both airborne and spaceborne platforms [15]. The SAR data acquisition process is portrayed in Figure 1.7, while four examples of SAR images, acquired respectively by Sentinel-1A, RADARSAT-2, COSMO-SkyMed, and

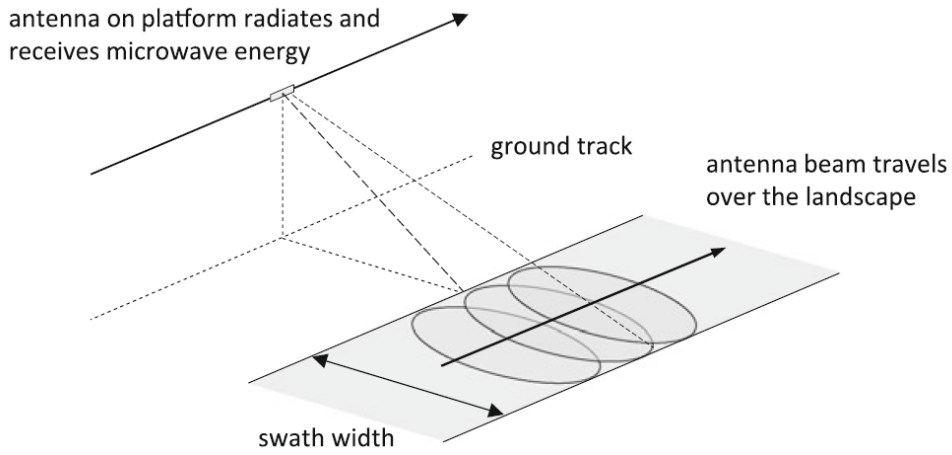


Figure 1.7: Synthetic aperture radar imaging. As the antenna beam travels over the features on the ground, the pulses transmitted from the platform to the ground generate many echoes that, once received, are processed to generate a very high resolution image of such features [1].

PALSAR-2, are shown in Figure 1.8.

It is worth noting that a radar system for EO operates regardless of Sun illumination, because it makes use of its own source of transmitted energy, and that the resulting data are almost insensitive to cloud cover and atmospheric conditions [1]. Therefore, unlike passive instruments, radar sensors for EO provide day-and-night and all-weather acquisition capability, thus complementing the properties of passive multispectral imagery. This complementarity between the physical natures and properties of the two typologies of remote sensing data is among the main reasons explaining the potential and relevance of their joint use within data fusion schemes.

In addition to the basic single-frequency and single-polarization mode, other configurations of radar EO also exist. Polarimetric SAR (PolSAR) collects (usually complex-valued) measurements associated with multiple polarizations simultaneously. Interferometric SAR (InSAR) exploits measurements of the phase of the radar return (and not only of its power) to extract 3D information on the observed surface. Differential InSAR (DInSAR) further extends InSAR to map slow movements of the surface (e.g., due to seismic phenomena). SAR tomography uses measurements taken from differ-

ent altitudes (e.g., different orbits) to characterize the vertical structures of the targets. Multifrequency SAR uses multiple antennas on-board the same platform to collect data at multiple carrier frequencies. Bistatic SAR uses distinct antennas for transmittance and reception to investigate the scattering behavior in multiple directions.

1.3 The Role of Resolution

The word "resolution" intuitively refers to the precision with which a given instrument captures details on the observed surface. More precisely, there are multiple meanings for the term "resolution", each one defined by the specific domain that is referred to.

First, spatial resolution is the size of the smallest spatial detail that can be distinguished in a remote sensing image [1]. It is related to the size of the ground area associated with a pixel. However, it can be slightly coarser because of the blurring effects that occur within the acquisition chain. In the case of a passive sensor, the spatial resolution depends on the sensor optics and on the altitude of the platform. In particular, the portion of the ground (in the across-track direction) that is observed while the platform travels his orbit is determined by the field of view of the sensor, while the instantaneous field of view determines the smallest detail that the sensor can capture, hence the spatial resolution Figure 1.9. In the case of a SAR instrument, it is possible to prove that the spatial resolution is related to the chirp processing and, within SAR technology, the resolution is also independent on the platform altitude [1].

The spatial resolutions of current satellite sensors for civil applications are approximately a few kilometers in the case of geostationary sensors (e.g., 3 km for the TIR bands of the SEVIRI sensor), a few tens of meters in the case of moderate resolution sensors (e.g., the Landsat series of satellite missions), and up to 30 cm - 1 m with recent very high resolution (VHR) near polar sensors (e.g., WorldView-2 and -3, Pléiades, COSMO-SkyMed, and TerraSAR-X). Spatial resolutions up to a few centimeters can usually be obtained using airborne acquisitions. For example, Figure 1.10 displays portions of six remote sensing images with the same size in pixels (400×400 pixels) and with very different spatial resolutions (3 km, 500 m, 30 m, 10 m, 2 m, and 5 cm). The difference in the spatial details that can be appreciated

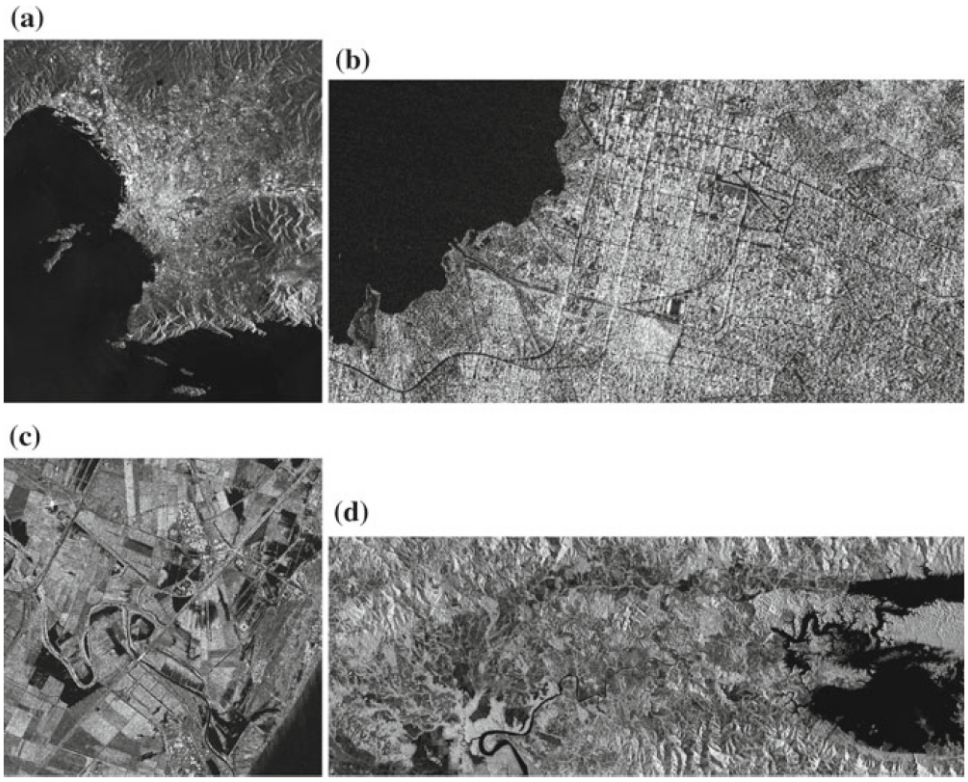


Figure 1.8: Examples of SAR images: (a) Sentinel-1A image acquired over Marseille, France (2100 × 2400 pixels); (b) RADARSAT-2 image acquired over Port-au-Prince, Haiti (1536 × 781 pixels); (c) COSMO-SkyMed (COnstellation of small Satellites for Mediterranean basin Observation) image acquired over the same area of Figure 1.6 shortly after a flood (2000 × 2000 pixels); and (d) PALSAR-2 (Phased Array type L-band Synthetic Aperture Radar 2, on-board the Advanced Land Observing Satellite 2, ALOS-2) image of a vegetated area and of Lake Bayano in Panama (9228 × 3471 pixels)

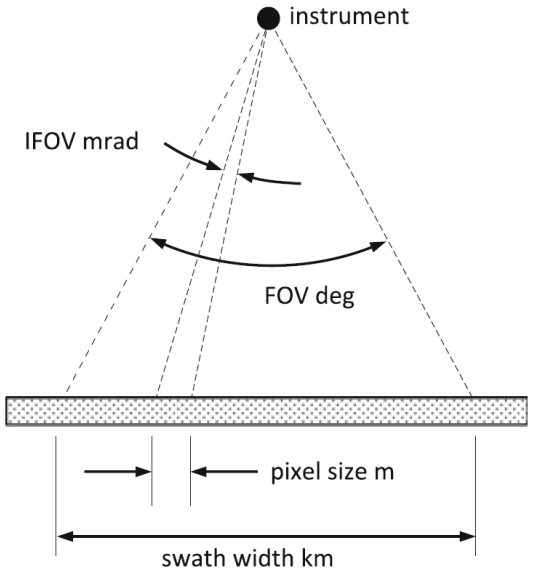
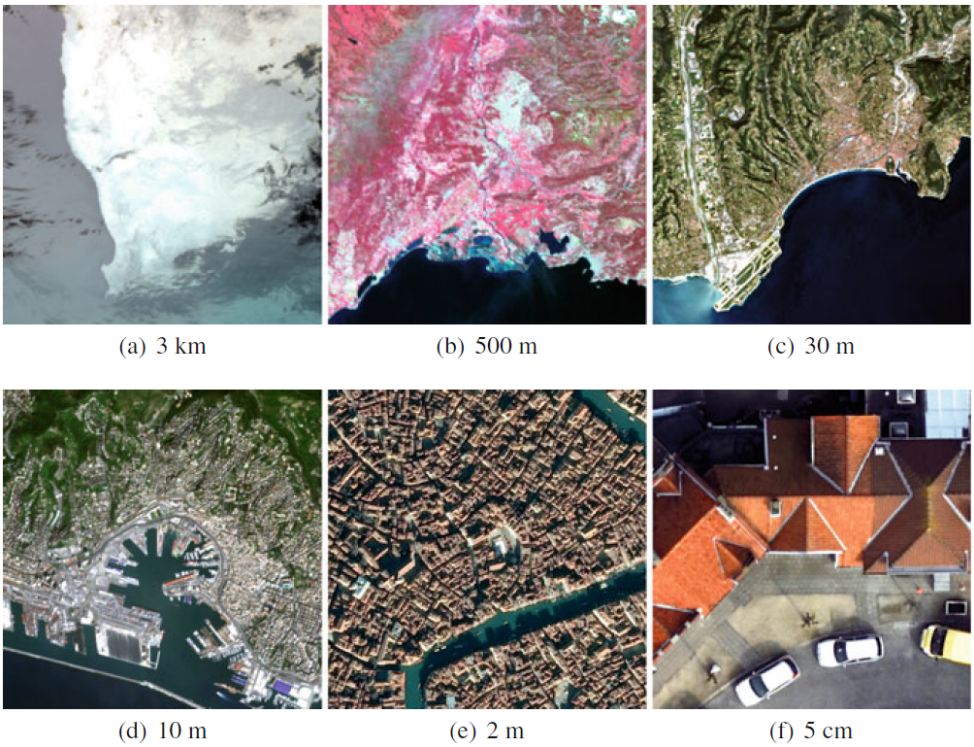


Figure 1.9: Definition of spatial resolution in remote sensing imagery [1].

in these images is visually evident.

The temporal resolution of a spaceborne sensor is the frequency with which a given ground area is repetitively observed. It is generally expressed in terms of the revisit time, i.e., the time between two consecutive satellite overpasses [1]. Typical values range from a few tens of minutes for geostationary sensors (e.g., approximately 15 min for MSG) to a few days or weeks for near polar sensors (e.g., one day for the Visible Infrared Imaging Radiometer Suite, VIIRS, and 16 days for Landsat-8). The use of multiple satellites in a constellation favors shorter revisit time (e.g., up to 12 hours for the COSMO-SkyMed constellation composed of four satellites). It is also worth noting that current near polar sensors often exhibit a pointing functionality, i.e., their observation directions can be steered upon agreement with the agency or company in charge of mission operations. This allows more frequent observations to be obtained on a given area but could sometimes make revisit times no more periodical and less predictable.

The spectral resolution is associated with passive multispectral sensors and is the precision with which the incoming radiation is sampled along the electromagnetic spectrum. It is usually expressed in terms of the number



(a) 3 km (b) 500 m (c) 30 m

(d) 10 m (e) 2 m (f) 5 cm

Figure 1.10: Color composites of channels from six multispectral images of size equal to 400 × 400 pixels: (a) SEVIRI image of South Africa; (b) Sentinel-3 Sea and Land Surface Temperature Radiometer (SLSTR) image of the south coast of France; (c) Landsat-8 Operational Land Imager (OLI) image of Nice, France; (d) Sentinel-2 image of Genoa, Italy; (e) Pléiades image of Venice, Italy; and (f) image collected by an airborne sensor over Zeebruges, Belgium. The spatial resolution is indicated below each image.

of channels of the sensor and of the widths of the corresponding wavelength ranges. Current sensors for civil applications range from a few bands of moderate width (70 - 100 nm each) to the case of hyperspectral sensors with a few hundreds narrow bands (2 - 10 nm each). For reasons associated with signal-to-noise ratio, a trade-off usually exists between the spectral and the spatial resolutions of a given passive sensor [11]. Therefore, several current satellite passive systems carry both a multispectral sensor, with several bands in the visible and NIR range, and an additional panchromatic sensor, which has obviously poorer spectral resolution but achieves finer spatial resolution than the multispectral bands.

Finally, radiometric resolution is related to the precision with which differences in the considered physical quantities can be appreciated and measured in the recorded image [1]. It is related to the signal-to-noise ratio of the sensor [11] and to the digitization process that is included in the acquisition chain. The intensity of a pixel in a digital image (or in each band of a multispectral digital image) is encoded with a finite number of bits, which correspond to a finite number of levels in a predefined discrete set (named quantization levels in the signal processing literature and often digital numbers in the remote sensing literature). The radiometric resolution of a sensor is generally expressed in terms of the number of bits that are used to encode each quantized intensity and are associated with each pixel (number of bits per pixel, bpp, sometimes also named bit depth). As shown in Figure 1.11, typical values range from 8 bpp (256 levels) to 12 bpp (4096 levels) and 16 bpp (65536 levels).

1.4 Methodological Contributions of the Thesis

This introductory chapter is meant to introduce the reader to the concepts underlying the remote sensing field, with details on the different types of sensors, acquisition modalities, measured physical quantities, and platforms carrying the remote sensors. Such a variety of technology highlights the need for automatic methods able to effectively exploit all the available information. Indeed, in the last years, lot of effort has been put in the design and development of advanced data fusion methods able to extract and make use of all the information available from as many information sources as possible.

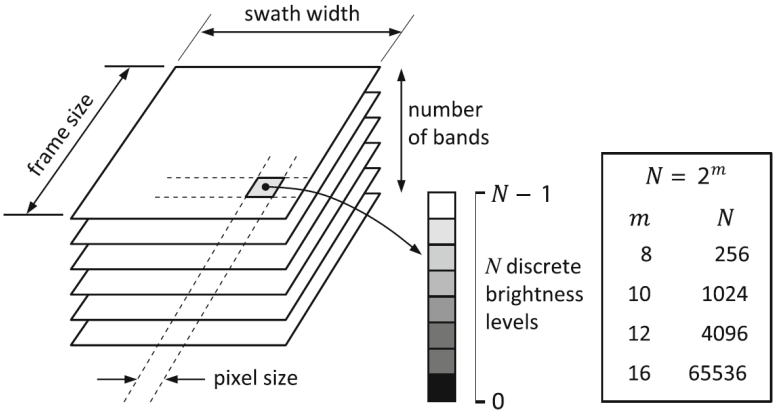


Figure 1.11: Definition of radiometric resolution in remote sensing imagery [1].

More specifically, in the broader context of data fusion, image registration plays a major role as accurate registration algorithms are essential in supporting Earth and planetary scientists [16]. Indeed, through image registration it is possible to bring two or more digital images into precise alignment for analysis and comparison. The main reason for the increased significance of image registration in remote sensing is the operational involvement in many important applications including, for example, the management of natural disasters, the assessment of climate changes, the management of natural resources, and the preservation of the environment. All such applications involve the monitoring of the Earth's surface over time and using as many information sources as possible. Furthermore, there is an increasing availability of images with different characteristics, thanks to shorter revisiting times of satellites, increased flexibility of use (different acquisition modalities) and the evolution of sensor technologies. Therefore, being able to simultaneously process different data for information extraction and fusion has become paramount in remote sensing. This includes the comparison of newly acquired images with previous images taken with different sensors or with different acquisition modalities or geometric configurations. The remote images can, therefore, be multitemporal (taken at different dates), multisource (derived from multiple sensors), multimode (obtained with different acquisition modalities), or stereo-images (taken from different viewpoints).

In the present thesis, this framework is addressed by operating in the context of machine learning and pattern recognition methodologies and by

proposing several innovative data fusion and image registration techniques aimed at solving specific problems of information extraction from remote-sensing images while making use of multi-source data. Furthermore, such innovative methodologies are applied to different typologies of data, ranging from planetary images (including images of Mars and the Moon), to images of the Earth surface, and taking into consideration both passive and active sensors. Moreover, the proposed methodologies are designed for supporting different applications, such as planetary science studies, the study of the climate change and how it is correlated with the changes in the Earth surface, and the identification of changes in a given area, possibly due to natural disasters or changes due to human intervention over the years.

In the context of planetary science studies, because of the large variety of planetary sensors and spacecraft already collecting data and with many new and improved sensors being planned for future missions, there is a strong need for integrating numerous multimodal image sources and, as a consequence, accurate and robust registration algorithms are required. Leveraging on the large variety of sensors and spacecraft already collecting planetary data, and based on the variety of new and improved sensors being planned for future missions, Chapter 2 deals with the need for automatic methods for the integration of such a large and heterogeneous amount of data. Indeed, the goal of the chapter is to propose and analyze a novel method for planetary image registration. The novel solution is based on a two-step registration process. The first step is based on matching a set of spatial features (i.e., the craters) extracted from the input images using a novel method based on the stochastic geometry modelling capabilities of marked point processes (MPP). The second step is based on an area-based information-theoretic functional optimized via simulated annealing, generalized pattern search, and genetic algorithms, which are computationally heavier, but restricted to a subset of the general registration problem thanks to the results achieved by the previous step.

While Chapter 2 focuses on the data fusion problem in the context of single-sensor data, Chapter 3 and Chapter 4 move the focus to those problems where the input data have been collected by different sensors (e.g., optical and radar imagery). More in details, the two chapters propose two novel solutions for the multisensor image registration problem. Chapter 3 proposes the use of the domain adaptation capabilities of the conditional generative adversarial networks to move the multisenor registration problem to the single-sensor

case, thus allowing the use of simpler and less computationally demanding similarity metrics. On the contrary, Chapter 4 proposes a solution for the same problem in a large-scale scenario. In this case, the goal is the design of a method that is accurate and robust, but that also requires few resources from the computational viewpoint to favor application to large-scale imagery in a climate-change application.

Finally, Chapter 5 deals with the case of multiresolution fusion and with another application in the context of Markov random field and machine learning methodologies applied to remotely sensed images. In particular, the chapter proposes a novel unsupervised change detection method that is able to cope with multimodality and multiresolution SAR imagery acquired at different times. From an application-oriented viewpoint, the method takes advantage of the multiresolution and multimodality acquisition capabilities of current satellite SAR missions. From a methodological perspective, given a pair of multiresolution and multimodal SAR images, the method is able to iteratively compute a change map at the finest resolution available in the input dataset. The method is based on Markovian probabilistic graphical models, graph cuts, linear mixtures, generalized Gaussian distributions, Gram–Charlier approximations, maximum likelihood and minimum mean squared error estimation. The statistics of the data at the various spatial scales is modelled through appropriate generalized Gaussian distributions and by iteratively estimating a set of virtual images that are defined on the pixel grid at the finest resolution and represent the data that would have been collected if all the sensors could work at that resolution. A Markov random field framework is adopted to address the detection problem by defining an appropriate multimodal energy function that can be minimized using graph cuts or belied propagation methods.

On a broader perspective and with an eye on the future trends, the remote sensing community is currently giving primary attention to the application of machine learning and pattern recognition methodologies to big data scenarios. As a consequence, more and more interest is currently surrounding data fusion methodologies like the ones proposed in this thesis. Indeed, they allow not only to extract information from specific data types, but also to jointly exploit multiple data sources. The collection of radar and optical data by Sentinel-1 and Sentinel-2, which is made available by the European Space Agency, together with the Landsat archives by NASA and USGS, and the recent international contests organized by communities like

the Geoscience and Remote Sensing Society, are a clear example of such an interest. Novel data fusion methodologies based on machine learning and pattern recognition have a huge impact in this scenario. Indeed, they allow integrating complementary data sources in diverse applications and, due to the heterogeneous nature of their input data, are flexible enough to meet the requirements of the ever growing sets of diverse data that are currently available and that will be available in the future.