Development of an Innovative System for the Reconstruction of New Generation Satellite Images

Luca Lorenzi

N° d'ordre : 2012telb0242

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En habilitation conjointe avec l'Université de Rennes 1

Co-tutelle avec Université de Trento (Italie)

Ecole Doctorale – MATISSE

Development of an Innovative System for the Reconstruction of New Generation Satellite Images

---

Thèse de Doctorat

Mention : Traitement du signal et télécommunications

Présentée par **Luca Lorenzi**

Département : Image et Traitement de l'Information

Laboratoire : Lab-STICC

Directeur de thèse : Grégoire Mercier

Soutenue le 29 novembre 2012

Jury :

M. Jocelyn Chanussot, Prof., Grenoble Institute of Technology (Rapporteur / Président)
M. Farid Melgani, Prof., University of Trento (Directeur de thèse)
M. Grégoire Mercier, Prof., Telecom Bretagne (Directeur de thèse)
M. Paolo Gamba, Prof., University of Pavia (Examinateur)

This work has be done with the financial support of the Italian-French University

"Queste erano le cose da fare, senza tralasciare quelle"

"These were the things to do, without neglecting those"

Mt 23

A Virginia

Abstract

Remote sensing satellites have demonstrated to be a helpful instrument. Indeed, satellite images have been successfully exploited to deal with several applications including environmental monitoring and prevention of natural disasters. In the last years, the increasing of the availability of very high spatial resolution (VHR) remote sensing images resulted in new potentially relevant applications related to land cover control and environmental management. In particular, optical sensors may suffer from the presence of clouds and/or of shadows. This involves the problem of missing data, which may result in an important problem especially in the case of VHR images.

In this thesis, new methodologies of detection and reconstruction of missing data region in VHR images are proposed and applied on areas contaminated by the presence of clouds and/or shadows. In particular, the proposed methodological contributions include: i) a multiresolution inpainting strategy to reconstruct cloud-contaminated images; ii) a new combination of radiometric information and spatial position information in two specific kernels to perform a better reconstruction of cloud-contaminated regions by adopting a support vector regression (SVR) method; iii) the exploitation of compressive sensing theory adopting three different strategies (orthogonal matching pursuit, basis pursuit and a genetic algorithm solution) for the reconstruction of cloud-contaminated images; iv) a complete processing chain which exploits a support vector machine (SVM) classification and morphological filters for the detection and a linear regression for the reconstruction of specific shadow areas; and v) several evaluation criteria capable to assess the reconstructability of shadow areas. All of them are specifically developed to work with VHR images.

Experimental results conducted on real data are reported in order to show and confirm the validity of all the proposed methods. They all suggest that, despite the complexity of the problems, it is possible to recover in a good way missing areas obscured by clouds or shadows.

Keywords

Cloud reconstruction, shadow reconstruction, very high resolution (VHR) images, remote sensing, pattern recognition, inpainting, support vector machine, support vector regression, compressive sensing.

Résumé

Les satellites de télédétection sont devenus incontournables pour la société civile. En effet, les images satellites ont été exploitées avec succès pour traiter plusieurs applications, notamment la surveillance de l'environnement et de la prévention des catastrophes naturelles. Dans les dernières années, l'augmentation de la disponibilité de très haute résolution spatiale (THR) d'images de télédétection abouti à de nouvelles applications potentiellement pertinentes liées au suivi d'utilisation des sols et à la gestion environnementale. Cependant, les capteurs optiques, en raison du fait qu'ils acquièrent directement la lumière réfléchie par le soleil, ils peuvent souffrir de la présence de nuages dans le ciel et / ou d'ombres sur la terre. Il s'agit du problème des données manquantes, qui induit un problème important et crucial, en particulier dans le cas des images THR, où l'augmentation des détails géométriques induit une grande perte d'informations.

Dans cette thèse, de nouvelles méthodologies de détection et de reconstruction de la région contenant des données manquantes dans les images THR sont proposées et appliquées sur les zones contaminées par la présence de nuages et / ou d'ombres. En particulier, les contributions méthodologiques proposées comprennent: i) une stratégie multirésolution d'inpainting visant à reconstruire les images contaminées par des nuages ; ii) une nouvelle combinaison d'information radiométrique et des informations de position spatiale dans deux noyaux spécifiques pour effectuer une meilleure reconstitution des régions contaminés par les nuages en adoptant une régression par méthode a vecteurs supports (RMVS) ; iii) l'exploitation de la théorie de l'échantillonnage compressé avec trois stratégies différentes (orthogonal matching pursuit, basis pursuit et une solution d'échantillonnage compressé, basé sur un algorithme génétique) pour la reconstruction d'images contaminés par des nuages; iv) une chaîne de traitement complète qui utilise une méthode à vecteurs de supports (SVM) pour la classification et la détection des zones d'ombre, puis une régression linéaire pour la reconstruction de ces zones, et enfin v) plusieurs critères d'évaluation promptes à évaluer la performance de reconstruction des zones d'ombre.

Toutes ces méthodes ont été spécialement développées pour fonctionner avec des images très haute résolution. Les résultats expérimentaux menés sur des données réelles sont présentés afin de montrer et de confirmer la validité de toutes les méthodes proposées. Ils suggèrent que, malgré la complexité des problèmes, il est possible de récupérer de façon acceptable les zones manquantes masquées par les nuages ou rendues erronées les ombres.

Mots-clés

Reconstruction de données manquantes, reconstruction de données erronées, Images à très haute résolution (THR), télédétection, reconnaissance de formes, Inpainting, méthode à vecteurs de support, régression à vecteurs de support, échantillonnage compressé.

Astratto

I satelliti per il telerilevamento stanno dimostrando sempre più di essere uno strumento molto utile, infatti, le immagini satellitari sono di grande beneficio per diverse applicazioni, tra cui il monitoraggio ambientale e la prevenzione di disastri naturali. Negli ultimi anni, l'aumento della disponibilità di immagini telerilevate ad altissima risoluzione spaziale ha comportato nuove applicazioni potenzialmente rilevanti per il controllo della copertura terrestre, dell'uso del suolo terrestre e per la gestione ambientale. In particolare, i sensori ottici hanno fornito un contributo straordinario per questi scopi. Tuttavia, a causa del fatto che essi acquisiscono direttamente la luce riflessa del sole, possono soffrire della presenza di nuvole nel cielo e / o di ombre sulla terra. Questo fatto potrebbe portare ad avere dei dati mancanti, il quale causerebbe un problema significativo specialmente nel caso di immagini ad altissima risoluzione spaziale, dove la quantità di informazione è particolarmente importante.

In questa tesi, nuove metodologie di rilevazione e di ricostruzione di regioni mancanti in immagini ottiche ad altissima risoluzione spaziale sono proposte e applicate per risolvere il problema di aree contaminate dalla presenza di nuvole e / o di ombre in tali immagini. In particolare, i contributi metodologici proposti sono: i) una strategia di inpainting multirisoluzione per ricostruire immagini contaminate dalla presenza di nuvole; ii) una nuova combinazione di informazioni radiometriche e di posizionamento spaziale suddivisa in due specifici kernel per effettuare una migliore ricostruzione di regioni contaminate dalla presenza di nuvole attraverso l'adozione di una regressione che sfrutta macchine a vettori di supporto (Support vector regression, SVR), iii) lo sfruttamento della teoria denominata compressive sensing, adottando tre diverse strategie (orthogonal matching pursuit, basis pursuit ed una soluzione basata su di un algoritmo genetico) per la ricostruzione di immagini contaminate dalla presenza di nuvole; iv) una catena completa di elaborazione che sfrutta una macchina a vettori di supporto (support vector machine, SVM) e filtri morfologici, con l'obiettivo di rilevare e ricostruire (mediante una regressione di tipo lineare) specifiche zone d'ombra, v) la valutazione di diversi criteri in grado di determinare la ricostruibilità delle zone d'ombra.

I risultati sperimentali condotti su dati reali sono riportati al fine di dimostrare e di confermare la validità di tutti i metodi qui proposti. Questi suggeriscono che, nonostante la complessità dei problemi, è possibile recuperare con buona accuratezza le zone mancanti oscurate dalla presenza di nuvole e / o di ombre.

Parole chiave

Ricostruzione di nuvole, ricostruzione di ombre, immagini ad altissima risoluzione, telerilevamento, riconoscimento di pattern, inpainting, macchine a vettori di supporto, regressione a vettori di supporto, compressive sensing.

Thanks to - Ringraziamenti

Ringrazio

- quello lassù che mi ha permesso di arrivare fin qui;
- la mia Grande Famiglia […] che mi ha sempre sostenuto;
- Il mio Amore Virginia, da 10 anni al mio fianco;
- Assieme alla sua piccola famiglia (Carolina, Roberto e Stefano);
- والاعتزاز الفخر مشاعر بخالص له أتقدم ولي لدعمه جفريد السيد أستاذي، الشكر جزيل أشكر أن أود
- Je remercie mon directeur de thèse Grégoire Mercier pour ses grands et riches conseils;
- I miei compagni d'avventura, Edo (+ Chiara) e Thomas, di continuo disponibili;
- Tutti voi qui presenti e quelli che non sono potuti venire.

Contents

List of Tables

List of Figures

Glossary

ALS: airborne laser scanner

AVIRIS: airborne visible infrared imaging spectrometer

BRDF: bidirectional reflectance distribution function

BP: basis pursuit

CCA: canonical correlation analysis

CNES: Centre National d'Études Spatiales

CV: cross-validation

CV: coefficient of variation

CS: compressive sensing

GA: genetic algorithm

HCV: hue, chroma, and value

HIS: hue, intensity and saturation

HSV: hue, saturation, and value

IR: infrared

LIDAR: light detection and ranging

KDE: kernel density estimation

MAD: multivariate alteration detection

MODIS: moderate resolution imaging spectrometer

MSER: maximally stable extremal regions

NDVI: normalized difference vegetation index

NSVDI: normalized saturation-value difference index

OMP: orthogonal matching pursuit

PAN: panchromatic

PCA: principal component analysis

PSNR: peak signal-to-noise ratio

RADAR: radio detection and ranging

RBF: radial basis function

ROI: region of interest

RS: remote sensing

TR: training

TS: test

SAR: synthetic aperture radar

SE: structuring element

SSE: sum square error

SV: support vector

SVM: support vector machine

SVR: support vector regression

VHR: very high resolution

YIQ: luma, inphase, and quadrature

YCbCr: luma and chroma, blue and red components

1. Introduction and Thesis Overview

Abstract *– In this first chapter, we make an introduction to the problem of reconstruction of missing data in very high spatial resolution images. In a first step, we report a brief overview on the remote sensing field and the general context in which the thesis is positioned. In a second step, the specific problems faced in the following chapters are introduced. Finally, we describe the proposed solutions and an overview of the thesis structure and organization.*

1.1. Context

Remote sensing (RS) is the acquisition of information (spectral, spatial, and temporal) of an object, by the use of recording sensing devices without coming into contact with the object. One of the most common recording sensing device is the artificial satellite, which can be equipped with 1) a passive or 2) an active sensor [1]. Passive sensors detect the radiation that is emitted or reflected by a region or a target of interest. The most common source of radiation is the sunlight, which is reflected by the objects and afterward measured by passive sensors such as infrared sensors and radiometers. The main advantages of passive sensors are a highest spatial resolution and the possibility to obtain unique fingerprints (also known as spectral signature) of the scanned material (exploiting the multispectral or hyperspectral bands). Active sensors emit a specific signal in order to obtain a backscattered energy from the observed target. The discrimination between objects and/or areas depends on the backscattered radiation. Examples of active sensors are the SAR and the LIDAR. This last measures the time delay between the emission and the return of the signal, measuring several properties like the location, the elevation, the speed and the direction of the target. The major advantages of active sensors, include the capability to obtain measurements anytime, regardless of the time of day or season and robust to the atmospheric conditions [2]-[3]. In this dissertation, we consider remote sensing systems for Earth observation, coming from satellites which mount passive sensors capable to acquire multispectral and very high spatial resolution (VHR) images.

Before the advent of VHR optical satellites, it was only possible to acquire images with spatial resolution of the order of several meters (e.g., *Landsat 7* with a resolution of 15 meters). The new generation of very high resolution optical satellites, like *QuickBird*, *IKONOS-2*, *Pléiades*, *WorldView-1* and *-2*, and *GeoEye-1* and *-2*, opens novel prospective thanks to the increased amount of spatial information they convey [4]-[5]. These new multispectral sensors are capable to register very important information for many potential applications related to environmental monitoring and land control and management (see an example of VHR image and its spectral signature in Figure 1.1).

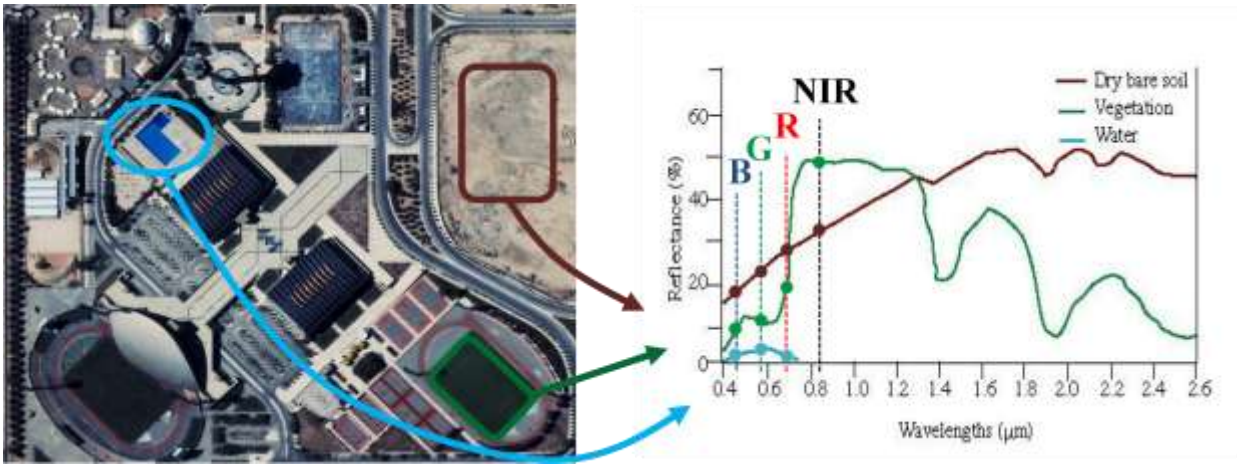


Figure 1.1. Example of an IKONOS-2 VHR image and related spectral bands (blue, green, red and near infrared).

New methodologies are however needed to analyze in an efficient way this new generation of data (e.g., classification, segmentation, edge detection, tree species identification, etc… [6]). A typical pattern recognition scheme is illustrated in Figure 1.2, where we want 1) to acquire the image with a proper sensor; 2) to pre-process the image in order to correct the problems which can occur during the acquisition system; 3) to extract some features, and at the same time select the best ones, in order to improve the work of the following step; 4) to classify in different thematic classes the acquired image.

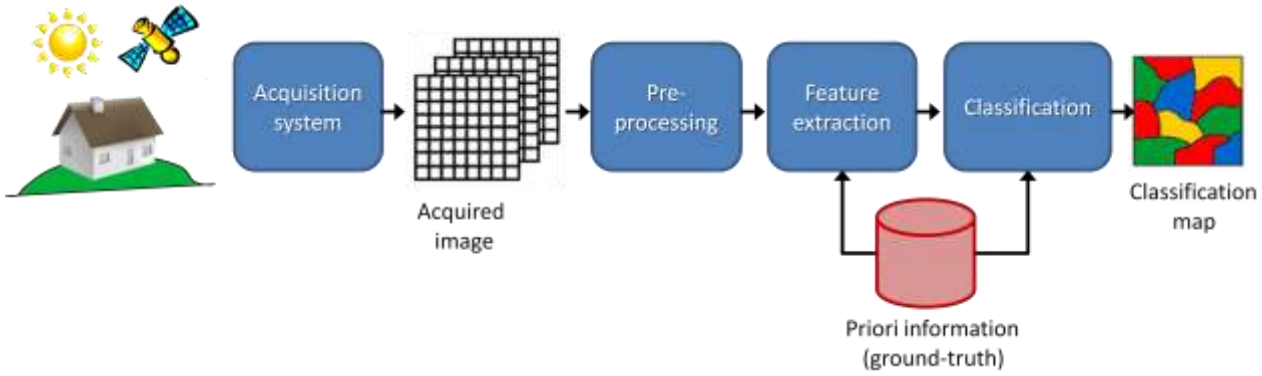


Figure 1.2. Example of a pattern recognition processing.

Potentially, with the advent of VHR images it is possible to obtain richer classification maps, containing very detailed classes. Unfortunately, passive sensors in general and VHR in particular suffer from the problem of atmospheric conditions, in particular from the presence of clouds and/or shadow covers [7]. As a consequence, exploitability of these images is affected by completely or partially missing data.

Figure 1.3. General block diagram of an image reconstruction process.

Despite the obscurations due to the presence of clouds and shadows have different origins, they can be considered as similar problems, but they require different implementations of the above described three principal steps. Nevertheless, for both cases, to reconstruct missing areas the main steps are (see Figure 1.3):

1. Image pre-processing: the first step is particularly important for it includes operations such as co-registration and calibration of the image.
2. Missing area detection: the scope of this second task is to identify the location of the missing areas. To obtain a precise, automatic and unsupervised identification of these regions, detection could exploit the statistical distributions, the correlations and the geometries of these areas. Besides the spatial and spectral information, also the temporal could be evaluated if available. Indeed considering different acquisitions on the same area

but at different times will hopefully help the detection thanks to the non-stationary nature of the cloud and shadow contributions. Since clouds and shadows refer to different contamination operations, for each of them, appropriate detection solutions must be considered.

3. Missing area reconstruction: after detecting the position of the missing areas, the goal of this last step is to restore them. Similarly to the previous step, reconstructing clouds or shadows must be handled differently. Cloud reconstruction appears a priori more complex than reconstructing shaded area, because shadows only partially affect the obscured information.

In the next paragraph we will describe in more detail the problems that each of these steps potentially convey.

1.2. Problems

Passive sensors are limited mainly by their sensitivity on the atmospheric condition during the acquisition of the image. The acquired images are frequently subject to the presence of clouds and/or of shadows. The former are affected by a complete missing data and may cover also big regions, depending on the season or on the geographical position of the acquisition (see Figure 1.4(a)). In the latter case, we have only a partial missing of data, and shadows are present in almost all the images, particularly in urban areas where there are larger changes in surface elevation (due to the presence of buildings, bridges, towers, etc) and consequently longer shadows (see Figure 1.4(b)). Depending on the application clouds and shadows can be viewed: 1) as a source of information for the evaluation of important parameters (e.g., in the case of clouds, the cloud liquid water in meteorological forecast and hydrological studies [8]-[9], whereas for the case of shadows, the building position and their height [10]) or 2) as a source of contamination which strongly affects the quality of the image and returns partially useless information. This last case is the subject on which we will focus our attention.

(a) (b)

Figure 1.4. Example of (a) completely obscuration by a cloud and (b) a partially obscuration by a shadow.

In both cases, clouds or shadows, the detection and the process to mask the undesired areas of missing data represents one of the first step to cope with. Clouds detection is in general based on the fact that clouds have a lower temperature and thus their spectral signature is lower in the infrared (IR) frequency. Moreover they are brighter, indeed they reflect almost all the sun illumination. In the literature, different strategies were proposed. Their general idea is to reduce the task to a binary classification problem, discerning between clouds and free-cloud regions [11]-[13]. For the detection of shadows the literature reports mainly two approaches: model-based and shadow properties-based approaches. The former needs *a priori* information about the scenario and the sensor. However, usually such knowledge is not available. Therefore most of the shadow detection algorithms are based on their properties, principally based on the fact that shaded areas present lower brightness, higher saturation and greater hue values (see Phong model [14]). In early works, shadow properties-based approaches attempt to detect shadows using a space color transformation and an automatic threshold estimator [15]. In a comparative work [16], several invariant color spaces comprising HIS, HSV, HCV, YIQ and YCbCr, were analyzed to detect shadows. Based on the results obtained in this study, a better shadow-detection approach was developed, using a novel successive thresholding scheme [17]. In addition there exist other algorithms which just integrate to the processing an extra feature image, to help to better discriminate shaded areas (e.g.: NDVI normalized difference vegetation index [18], NSVDI normalized saturation-value difference index [19], MSER maximally stable extremal regions [20]). Another technique applies the principal component analysis (PCA) transformation to isolate the luminance component, exploiting the image multidimensionality [21].

The restoration of digital images has been deeply studied in several applications mainly because of its theoretical and practical importance in different fields, such as radioastronomy, biomedical engineering, and machine vision [22]. Coming now to the RS field, also here a certain attention has been paid for the reconstruction of digital images and in different problems, like acquisition blur and geometric distortions [23], phase distortions [24], resampling problems [25], or problems related to applications like buried object detection [26]. But for the cases of clouds and shadows restoration relatively few works were presented in the literature, especially for the case of VHR images.

Focusing on the problem of clouds, in the last years, different works have been presented and are mainly intended for low or medium spatial resolution images and are based on the assumption that the temporal signature of a given pixel is contaminated by residual effects caused by imperfect sensing of the target or by spatially autocorrelated noise due to atmospheric attenuation. They make use of principles such as the data substitution principle [27] or temporal prediction strategies [28]. Another approach aims at removing a particular kind of clouds, the cirrus clouds, acquired from the moderate resolution imaging spectrometer (MODIS) or from the airborne visible infrared imaging spectrometer (AVIRIS) [29]. The authors found that the measurements acquired at the 1.38 μm band are essentially due to the bidirectional reflectance of cirrus cloud attenuated by the absorption of water vapor above this cirrus. They use this fact to correct and remove these attenuation effects. More recently other techniques can resort to the filling-in approach such as inpainting techniques, which aim at filling holes in digital images by propagating surrounding structures. An example is the inpainting method that allows removing clouds by means of the bandlet transform (a special

case of the wavelet transform) and the multiscale geometrical grouping. Its promising experimental results encourage the idea of exploiting inpainting techniques to reconstruct missing data in remote sensing imagery [30]. All these previous techniques present several drawbacks, for example, the sensor, the ground cover and the cloud typology dependence, the need of a very high temporal resolution and the higher methodological complexity.

Coming now to the shadow reconstruction, there exist essentially three different approaches: gamma correction, histogram matching and linear correlation [31]. In [32],the authors consider that the surface texture does not radically change when it is shaded. To remove shadows, the authors use a contextual texture analysis between a segment of shadow and its neighbors. Knowing the kind of surface under the shadow a local gamma transformation is used to restore the shaded area. On the contrary, in [33], after the detection of the shadows, the authors propose to adjust the hue, intensity and saturation values (HIS) in shadowed regions respectively according to the analogous values in the local surroundings around each shaded region, adopting the histogram matching method. In [34], the authors recover spectral information in shadow areas in an Ikonos image having the height data from the airborne laser scanner (ALS). Having evaluated a simulation of the possible shadow areas exploiting ALS data, the authors make use of this information to overlay and eliminate the real shadow, comparing the results obtained with two methods: gamma correction and linear correlation. In [35], the authors consider the fact that the restoration of shadows almost depend on the spectral signature of the spectral bands. So first the bands are thresholded in an independent way determining the optimal threshold values by visual inspection. Then a linear regression in each spectral band is carried out to correct the shadow effects. All these previous techniques adopt one of the most common solutions as a singular measurement. They can fail to restore obscured area if not adopted in a proper way; for example, if they obtain only one shadow restoration for the several thematic classes presented in the image.

Almost all of these reconstruction techniques may not work properly if used on images with a very high spatial resolution. The increment of the spatial resolution may be translated in a corresponding increase in the heterogeneity of the surfaces, in other words, in much more details. Accordingly, specific techniques should be developed in order to fully exploit all the potential conveyed by VHR optical images.

1.3. Thesis Objective, Solutions and Organization

As introduced in the previous subsection, missing data problem due to the presence of clouds and/or shadows represents a research field of great interest mainly because of its important implications. In this thesis work, we will focus on the reconstruction of image data, totally obscured by the presence of clouds or only partially obscured by the presence of shadows, by proposing new and innovative methodological solutions.

After this introductive section followed by a French version, the thesis is organized into six chapters. A first part, containing three sections, is devoted to the problem of cloud contamination, whether a second part, containing two sections, deals with the problem of shadow contamination.

In Chapter 3, the problem of missing area reconstruction due to the presence of clouds is faced. In particular, we focus on the reconstruction exploiting the inpainting methodology on a single date image. After a brief introduction of the region-based inpainting (RBI) algorithm developed by Criminisi *et al* [36], three different new strategies are described: 1) Feature extraction-based inpainting (FEBI), where additional textural features extracted from the original image are exploited; 2) inpainting with isometric transformation (IsoI), where we aim at increasing the amount of patch candidates and thus populating further the search space; and 3) multiresolution inpainting (MRI), where we progressively inject multiresolution information in order to obtain a better reconstruction.

In Chapter 4, a new strategy for support vector regression (SVR) in the context of multitemporal multispectral remote sensing images is proposed. In particular, we intend to improve the reconstruction process by integrating both radiometric and spatial position information. For each kind of information adopted in the regression, a specific kernel is selected and adapted; in more detail, we evaluate the performances of the combinations of three different typologies of kernels (e.g., linear-linear, linear-polynomial, etc.). Subsequently, their fusion is performed by a linear combination of the two resulting kernels. Support vector regression is applied to derive the prediction function.

In Chapter 5, three novel methods to solve the problem of the reconstruction of missing data due to the presence of clouds are proposed. Given a multitemporal dataset, missing measurements are recovered applying the compressive sensing (CS) theory in which cloud-free pixels are exploited. In more detail, we first adopt two of the most common CS methods which approximate the CS solution, namely the basis pursuit (BP) and the orthogonal matching pursuit (OMP). Furthermore, we propose an alternative CS solution, which exploits the search capabilities of genetic algorithms (GAs).

Coming now to the second part of the thesis, in Chapter 6, a new solution for the problem of the presence of shadows in very high resolution (VHR) is introduced. We propose a complete processing chain, which relies on various advanced image processing and pattern recognition tools. In more details, the detection of the shadow regions is made by a classification task, implemented by a support vector machine (SVM) approach, whereas the reconstruction is based on a linear regression method, capable to adjust the intensities of the shaded pixels accordingly to the corresponding non-shadow regions.

In Chapter 7, different criteria useful to help in understanding *a priori* if it is possible or not to reconstruct a specific shadow area are proposed. In particular, we assume that an ideal reconstructability criterion should not tolerate that an unreconstructable shadow area is assigned as reconstructable and, at the same time, should maximize the probability of detection of reconstructable areas. Several evaluation criteria working at the pixel and textural levels are presented, reported and discussed, in order to lead to the definition of a final global index based on the fusion of two single criteria which are the Kullback-Leibler divergence and the angular second-moment difference.

Finally, in the last chapter, concluding remarks on the proposed techniques are given, with some open issues and future research works.

This manuscript supposes that the reader has basic knowledge in the field of remote sensing, pattern recognition and image processing. On the contrary, we suggest her/him to consult the references available at the end of this Chapter, in order to obtain a better introduction on the basic concepts which will be adopted in the following chapters. Note that every chapter in this manuscript has been written so that they are as much self-contained as possible.

1.4. References

[1] Jian Guo Liu, and Philippa J. Mason, *Essential Image Processing for GIS and Remote Sensing*, Chichester: Wiley-Blackwell, 2009.

[2] Robert A. Schowengerdt, *Remote sensing: Models and methods for image processing*, 3rd ed., Tucson: Academic Press, 2007.

[3] John Robert Schott, *Remote sensing: The image chain approach*, 2nd ed., New York: Oxford University Press, 2007.

[4] R. Hellermann, "Growing demand for VHR satellite imagery for civilian use," in *Proc. International Commercial Remote Sensing Symposium*, Washington, USA, Mar. 2010.

[5] K. Jacobsen, "Very high resolution satellite images – competition to aerial images," in *Proc. Geospatial World Forum*, Hyderabad, India, Jan. 2011.

[6] A. P. Carleer, O. Debeir, and E. Wolff, "Assessment of very high spatial resolution satellite image segmentations," *J. of the American Society for Photogrammetry and Remote Sensing*, vol. 71, pp. 1285–1294, Nov. 2005.

[7] M. Neubert and G. Meinel, "Atmospheric and terrain correction of IKONOS imagery using ATCOR3," in *Proc. ISPRS*, Hannover, Germany, May 2005.

[8] J. R. Wang, P. Racette, M. E. Triesky, E. V. Browell, S. Ismail, and L. A. Chang, "Profiling of atmospheric water vapor with MIR and LASE," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1211–1219, Jun. 2002.

[9] B. G. Vasudevan, B. S. Gohil, and V. K. Agarwal, "Backpropagation neural-network-based retrieval of atmospheric water vapor and cloud liquid water from IRS-P4 MSMR," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 5, pp. 985–990, May 2004.

[10] T. Kim, T. Javzandulam and T-Y. Lee, "Semiautomatic reconstruction of building height and footprints from single satellite images," in *Proc. IGARSS*, Barcelona, Spain, Jul. 2007, pp. 4737–4741.

[11] L. L. Stowe, P. A. Davis, and E. P. McClain, "Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the advanced very high resolution radiometer," *J. Atmospheric and Oceanic Technology*, vol. 16, pp. 656–681, Jun. 1999.

[12] A. V. Di Vittorio, and W. J. Emery, "An automated, dynamic threshold cloud-masking algorithm for daytime AVHRR images over land," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1682–1694, Aug. 2002.

[13] F. Murtagh, D. Barreto, and J. Marcello, "Decision boundaries using Bayes factors: The case of cloud masks," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2952–2958, Dec. 2003.

[14] J. Liu, J. Yang, and T. Fang, "Color property analysis of remote sensing imagery," *Act Photonica Sinica*, vol. 38, no. 2, pp. 441–446, Feb. 2009.

[15] G. Finlayson, and S. Süsstrunk, "Optimization for Hue constant RGB sensors," in *Proc. IS&T/SID*, Portland, OR, USA, Apr. 2002, vol.10, pp. 343–348.

[16] V. Tsai, "A comparative study on shadow compensation of color aerial images in invariant color models," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no.6, pp. 1661–1671, Jun. 2006.

[17] K. L. Chung, Y. R. Lin, and Y. H. Huang, "Efficient shadow detection of color aerial images based on successive thresholding scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 671–682, Feb. 2009.

[18] D. Cai M. Li, Z. Bao, Z. Chen, W. Wei, and H. Zhang, "Study on shadow detection method on high resolution remote sensing image based on HIS space transformation and NDVI index," in *Proc. International Conference on Geoinformatics*, Beijing, China, Jun. 2010, pp.1–4.

[19] H. Ma, Q. Qin, and X. Shen, "Shadow segmentation and compensation in high resolution satellite images," in *Proc. IGARSS*, Boston, MA, USA, Jul. 2008, vol. 2, pp. 1036–1039.

[20] H.Y. Yu, J.G. Sun, L.N. Liu, Y.H. Wang, and Y.D. Wang, "MSER based shadow detection in high resolution remote sensing image," in *Proc. ICMLC*, Qingdao, China, Jul. 2010, pp. 780–783.

[21] S. Wang, and Y. Wang, "Shadow detection and compensation in high resolution satellite images based on retinex," in *Proc. ICIG*, Xi'an, China, Sep. 2009, pp. 209–212.

[22] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signals Process. Mag.*, vol. 14, no. 2, pp. 24–41, Mar. 1997.

[23] S. E. Reichenbach, D. E. Koehler, and D. W. Strelow, "Restoration and reconstruction of AVHRR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 997–1007, Jul. 1995.

[24] V. P. Bakalov and M. Y. Yerokhin, "Removal of uncontrollable phase distortions in synthetic aperture radar signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1298–1302, May 2000.

[25] S. E. Reichenbach and J. Li, "Restoration and reconstruction from overlapping images for multi-image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 4, pp. 769–780, Apr. 2001.

[26] Z. Wu and C. Liu, "An image reconstruction method using GPR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 327–334, Jan. 1999.

[27] S. C. Liew, M. Li, and L. K. Kwoh, "Automated production of cloud-free and cloud-shadow image mosaics from cloudy satellite imagery," in *Proc. ISPRS*, Istanbul, Turkey, Jul. 2004, pp. 523–530.

[28] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens*., vol. 44, no. 2, pp. 442–455, Feb. 2006.

[29] B.-C. Gao, P. Yang, W. Han, R.-R. Li, and W. J. Wiscombe, "An algorithm using visible and 1.38-μm channels to retrieve cirrus cloud reflectances from aircraft and satellite data," *IEEE Trans. Geosci. Remote Sens*., vol. 40, no. 8, pp. 1659–1668, Nov. 2002.

[30] A. Maalouf, P. Carré, B. Augereau and C. F. Maloigne, "A bandelet-based inpainting technique for clouds removal from remotely sensed images," *IEEE Trans. Geosci. Remote Sens*., vol. 47, no. 7, pp. 2363–2371, Jul. 2009.

[31] P. Sarabandi, F. Yamazaki, M. Matsuoka, and A. Kiremidjian, "Shadow detection and radiometric restoration in satellite high resolution images," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 6, pp. 3744–3747.

[32] A. Massalabi, H. Dong-Chen, G.B. Benie and E. Beaudry, "Detecting information under and from shadow in panchromatic IKONOS images of the city of Sherbrooke," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 3, pp. 2000–2004.

[33] J. Su, X. Lin, and D. Liu, "An automatic shadow detection and compensation method for remote sensed color images," in *Proc. ICSP*, Beijing, China, Nov. 2006, vol. 2.

[34] T. Nakajima, G. Tao, and Y. Yasuoka, "Simulated recovery of information in shadow areas on IKONOS image by combining ALS data," in *Proc. ACRS*, Kathmandu, Nepal, 2002.

[35] F. Yamazaki, W. Liu, and M. Takasaki, "Characteristic of shadow and removal of its effects for remote sensing imagery," in *Proc. IGARSS*, Cape Town, South Africa, Jul. 2009, vol. 4, pp.426–429.

[36] A. Crimisi, P. Perez and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. on Image Process*., vol. 13, no. 9, pp. 1–14, Sep. 2004.

2. Résumé Long en Français

2.1. Contexte

La télédétection est l'acquisition d'informations (spectrale, spatiale et temporelle) d'un objet, par l'utilisation d'appareils d'enregistrement de détection à distance, sans entrer en contact avec l'objet. Un des dispositifs d'enregistrement de la détection les plus fréquents provient de satellite artificiel, qui peut être équipé de 1) un capteur passif ou 2) un actif [1]. Les capteurs passifs peuvent détecter le rayonnement qui est émis ou réfléchi par une région ou une cible d'intérêt. La source la plus commune de rayonnement est la lumière du soleil, qui est réfléchie par les objets et ensuite mesurée par des capteurs passifs tels que des capteurs infrarouges et des radiomètres. Les principaux avantages de ces capteurs passifs sont une haute résolution spatiale et la possibilité d'obtenir des empreintes digitales uniques (également connu sous le nom signature spectrale) de la matière scannée (en exploitant les caractéristiques multispectrales ou hyperspectrales). Les capteurs actifs émettent un signal spécifique afin d'obtenir une énergie rétrodiffusée par la cible observée. La discrimination entre les objets et / ou de zones dépend du rayonnement rétrodiffusé. Les exemples de capteurs actifs sont les SAR et le LIDAR. Cette dernière mesure du délai entre l'émission et le retour du signal, la mesure de plusieurs propriétés telles que l'emplacement, l'altitude, la vitesse et la direction de la cible. Les principaux avantages de capteurs actifs sont notamment la capacité d'obtenir des mesures à tout moment, quel que soit le moment de la journée ou de la saison et d'être robuste aux conditions atmosphériques [2]-[3]. Dans cette dissertation, nous considérons des systèmes de télédétection pour l'observation de la Terre, venant de satellites passifs capables d'acquérir des images multispectrales à très haute résolution spatiale.

Avant l'ère des satellites optiques à très haute résolution, les précédents étaient capables d'acquérir des images avec une résolution spatiale qui était de l'ordre de plusieurs mètres (par exemple, Landsat 7 a été capable d'acquérir une image dans une bande panchromatique avec une résolution de 15 mètres). La nouvelle génération de satellites optiques à très haute résolution, comme QuickBird, Ikonos-2, Pléiades, WorldView-1 et -2, et GeoEye-1 et -2, ouvre de nouveaux potentiels grâce à l'augmentation du montant de l'information spatiale, de l'ordre du mètre, qu'ils véhiculent [4]-[5]. Ces nouveaux capteurs multispectraux sont capables d'enregistrer des informations très importantes pour de nombreuses applications potentielles liées à la surveillance et gestion de l'environnement, ainsi qu'à la maîtrise pression immobilière (voir un exemple d'image a très haute résolution et sa signature spectrale de la figure 2.1).

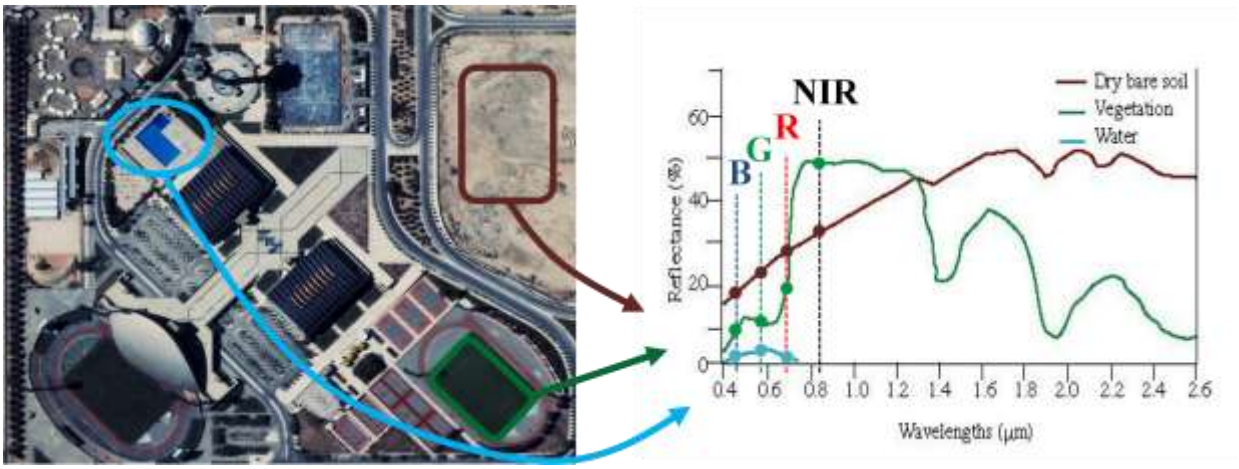


Figure 2.1. Exemple d'une image IKONOS-2 et liées bandes spectrales (bleu, vert, rouge et proche infrarouge).

De nouvelles méthodes sont toutefois nécessaires pour analyser d'une manière efficace cette nouvelle génération de données (par exemple, la classification, segmentation, détection de contours, arbre d'identification des espèces, etc. [6]). Un système de reconnaissance typique est illustré à la figure 2.2, où nous voulons 1) pour acquérir l'image, un capteur approprié, 2) pré-traiter de l'image afin de corriger les problèmes qui peuvent survenir lors de l'acquisition; 3) extraire certaines caractéristiques, et en même temps, sélectionner les meilleurs d'entre elles, afin d'alléger les travaux de l'étape suivante; 4) classer dans différentes classes thématiques de l'image ainsi acquise.

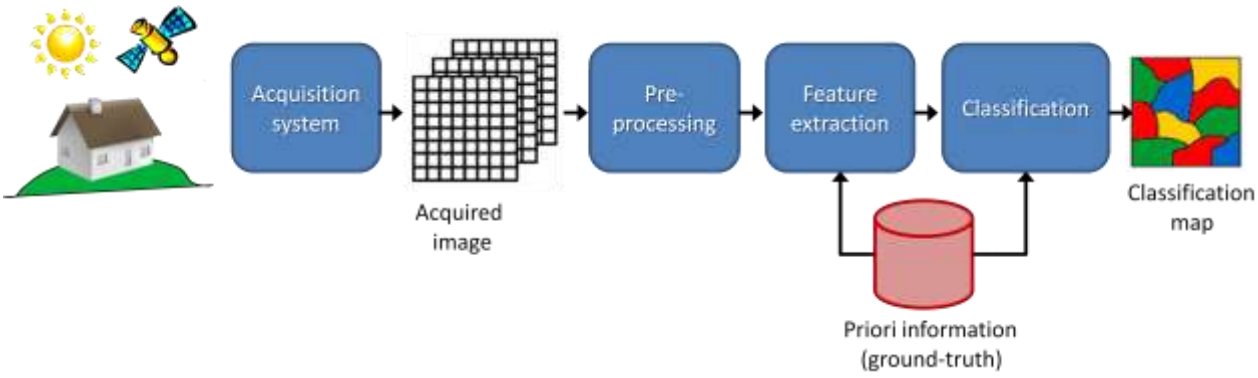


Figure 2.2. Exemple d'un traitement de reconnaissance de formes.

Potentiellement, avec l'avènement des images a très haute résolution, il est possible d'obtenir des cartes de classification les plus riches, contenant des classes très détaillées. Malheureusement, les capteurs passifs en général et en particulier a très haute résolution souffrent des perturbations induitent pas les conditions atmosphériques, en particulier par la présence de nuages et / ou de la couverture nuageuse [7]. En conséquence, l'exploitabilité de ces images est affectée par tout ou partie des données manquantes.

Figure 2.3. Schéma bloc général d'un processus de reconstruction de l'image.

Bien que les obstructions dues à la présence de nuages et les ombres aient des origines différentes, elles peuvent être considérées comme des problèmes similaires, mais elles nécessitent différentes implémentations des trois principales étapes décrites ci-dessus. Cependant, dans les deux cas, reconstituer les parties manquantes représente la principale étape (voir Figure 2.3):

1. pré-traitement de l'image: la première étape est particulièrement importante car elle comprend des opérations telles que la co-registration et l'étalonnage de l'image.
2. détection de la zone manquant: la portée de cette seconde tâche consiste à identifier l'emplacement des zones manquantes. Pour obtenir une identification précise, automatique et sans surveillance de ces régions, la détection pourrait exploiter les distributions statistiques, les corrélations et les géométries de ces régions. Outre l'information spatiale et spectrale, l'information temporelle pourrait aussi être évaluée si elle est disponible. Puisque les nuages et les ombres correspondent à des opérations de contaminations différentes, pour chacun d'eux, des solutions de détection appropriés doivent être pris en considération.
3. reconstruction de la zone manque: après la détection de la position des zones manquantes, le but de cette dernière étape consiste à les restaurer. Comme pour l'étape précédente, la reconstruction des nuages ou des ombres doit être traitée de façon spécifique. La reconstruction des zones cachées par les nuages semble plus complexe que la reconstruction des zones obscurcies ou cachées par les ombres a priori, car les ombres n'affectent que partiellement les informations initiales.

Dans le paragraphe suivant, nous allons décrire plus en détail les problèmes que chacune de ces étapes potentiellement véhiculent.

2.2. Problèmes

Les détecteurs passifs sont limités principalement par leur sensibilité aux conditions atmosphériques lors de l'acquisition de l'image. Les images acquises sont souvent affectées par la présence de nuages et / ou d'ombres. Les premières sont caractérisées par un manque de données complètes et peuvent également couvrir de grandes régions, en fonction de la saison ou de la position géographique de l'acquisition (voir la figure 2.4 (a)). Dans second cas, nous n'avons qu'un manque partiel des données et des ombres présentes dans presque toutes les images, en particulier dans les zones urbaines où il ya des grands changements d'élévation de surface (en raison de la présence de bâtiments, de ponts, tours, etc.) et par conséquent, des ombres plus longues (voir la figure 2.4 (b)). Selon les applications, les nuages et les ombres peuvent être considérées: 1) en tant que source d'informations pour l'évaluation des paramètres caractéristiques (par exemple, dans le

cas des nuages, l'eau liquide en nuage prévisions météorologiques et hydrologiques études [8]-[9], alors que pour le cas des ombres, la position et leur hauteur des bâtiments [10]) ou 2) comme source de contamination qui affectent fortement la qualité de l'image et envoie les informations partiellement inutiles. Ce dernier cas est le sujet sur lequel nous allons concentrer notre attention.

(a) (b)

Figure 2.4. Exemple d'images (a) complètement bouchée par un nuage et (b) un affectée partiellement par une ombre.

Dans les deux cas, les nuages ou les ombres, la détection et le processus de masquer les zones non souhaitées de données manquantes représentent une première étape considérée dans le traitement. La détection de nuages est en général basée sur le fait que les nuages ont une température plus basse et donc leur signature spectrale est plus faible dans l'infrarouge (IR). En outre, ils sont plus brillants, en effet, ils reflètent la quasi-totalité de l'éclairement. Dans la littérature, différentes stratégies ont été proposées. Leur idée générale est de réduire la tâche à un problème de classification binaire, discerner entre les nuages et les régions libre [11]-[13]. Pour la détection des ombres, la littérature rapporte principalement deux approches: approches à base de modèles et approches basées sur les propriétés des ombres. La première a besoin d'une information a priori sur la scène et le capteur. Cependant, le plus souvent de telles connaissances ne sont pas disponibles. C'est pourquoi la plupart des algorithmes de détection des ombres sont basés sur leurs propriétés intrinsèques, principalement basées sur le fait que les zones ombragées présentent une luminosité plus faible, de même pour la saturation et la teinte (voir modèle de Phong [14]). Dans les premières études, ont été proposées des approches basées sur les propriétés des ombres en utilisant une transformation de l'espace couleur et l'estimation automatique d'un seuil [15]. Dans une étude comparative [16], plusieurs espaces de couleur invariants comprenant HIS, HSV, le HCV, YIQ et YCbCr, ont été analysés pour détecter les ombres. Sur la base des résultats obtenus dans cette étude, une nouvelle approche a été proposée, en utilisant un système de seuillages successifs [17]. En outre, il existe d'autres algorithmes qui utilisent une image supplémentaire, pour aider à mieux distinguer les zones d'ombre (par exemple: l'indice végétation par différence normalisée NDVI [18], l'indice normalisé différence de saturation de valeur NSVDI [19], maximum stable régions extrêmes MSER [20]). Une autre technique utilise l'analyse en composantes principales (ACP) afin d'isoler la composante de luminance, en exploitant le caractère multidimensionnel de l'image [21].

La restauration des images numériques a été profondément étudiée dans plusieurs applications principalement en raison de son importance théorique et pratique dans différents domaines, tels que la radioastronomie, le génie biomédical, et la vision artificielle [22]. Venons-en maintenant au champ de la télédétection, là aussi une certaine attention a été accordée à la reconstruction des images numériques et des problèmes différents, comme le flou d'acquisition et de distorsions géométriques [23], la phase distorsions [24], les problèmes de ré-échantillonnage [25], ou problèmes liés à des applications telle la détection d'objets enfouis [26]. Mais pour le cas des nuages et des travaux de restauration des ombres, relativement peu d'études ont été présentées dans la littérature, encore moins dans le cas des images a très haute résolution.

Pour mettre l'accent sur le problème des nuages, dans ces dernières années, différents travaux ont été présentés et sont principalement destinés à la basse ou moyenne résolution spatiale des images. Ils sont fondés sur l'hypothèse que la signature temporelle d'un pixel donné est contaminée par des effets résiduels causés par une détection imparfaite ou par le bruit spatialement autocorrélées en raison de l'atténuation atmosphérique. Ils font usage de principes tels que le principe de substitution des données [27], ou des stratégies de prédiction temporelle [28]. Une autre approche vise à supprimer un type particulier de nuages, les nuages cirrus, acquis auprès du spectromètre MODIS ou AVIRIS [29]. Les auteurs ont constaté que les mesures acquises à la bande de 1,38 um sont essentiellement dues à la réflectance bidirectionnelle de cirrus atténuée par l'absorption de la vapeur d'eau au-dessus de ce cirrus. Ils utilisent ce fait pour corriger et éliminer ces effets d'atténuation. Plus récemment, d'autres techniques peuvent recourir à stratégies de reconstruction telles les techniques d'in-painting, qui visent à combler les trous dans les images numériques en propageant les structures environnantes. Un exemple est la méthode retouches qui permet d'enlever les nuages à l'aide de la transformée de bandelettes (un cas particulier de la transformée en ondelettes) et le regroupement multi-échelle géométrique. Ses résultats expérimentaux prometteurs ont encouragés l'idée d'exploiter les techniques de retouches pour reconstruire les données manquantes en télédétection [30]. Toutes ces techniques présentent plusieurs contraintes liées notamment aux capteurs, à la couverture du sol et à la dépendance typologie des nuages, à la nécessité d'une très haute résolution temporelle et une plus grande complexité méthodologique.

Venons-en maintenant à la reconstruction des zones affectées par de l'ombre. Il existe essentiellement trois approches différentes: la correction gamma, la correspondance de l'histogramme et la régression linéaire [31]. Dans [32], les auteurs considèrent que la texture de la surface ne change pas radicalement quand elle est à l'ombre. Pour supprimer les ombres, les auteurs utilisent une analyse de texture contextuelle entre un segment de l'ombre et ses voisins. Connaissant le type de surface à l'ombre, une transformation gamma locale est utilisée pour restaurer la zone ombragée. Au contraire, dans [33], après la détection de l'ombre, les auteurs proposent d'ajuster les valeurs de teinte, de saturation et d'intensité (HIS) dans les régions ombragées, respectivement, en fonction des valeurs analogues de l'environnement local autour de chaque région ombragée, en adoptant le procédé d'histogramme équivalent. Dans [34], les auteurs récupèrent l'information spectrale des zones d'ombre dans une image Ikonos et utilisent les données de hauteur d'un laser aéroporté (ALS). Après avoir évalué une simulation des zones d'ombre éventuelles à l'aide des données de l'ALS, les auteurs utilisent cette information pour superposer et éliminer l'ombre réelle,

en comparant les résultats obtenus avec deux méthodes: la correction gamma et la régression linéaire. Dans [35], les auteurs considèrent que la restauration des ombres dépendra de la signature spectrale des bandes spectrales. Les bandes sont d'abord seuillées de manière indépendante et la détermination des valeurs des seuils optimaux est effectuée par inspection visuelle. Puis une régression linéaire dans chaque bande spectrale est effectuée pour corriger les effets d'ombre. Toutes les techniques de l'état de l'art adoptent l'une de ces solutions, qui sont les plus courantes, comme une mesure singulière.

La quasi-totalité de ces techniques de reconstruction peuvent ne pas fonctionner correctement sur des images avec une très haute résolution spatiale. L'augmentation de la résolution spatiale s'accompagne d'une augmentation correspondante de l'hétérogénéité des surfaces, en d'autres termes, beaucoup plus de détails. Par conséquent, des techniques spécifiques doivent être développées afin d'exploiter pleinement tout le potentiel véhiculé par les images optiques à très haute résolution.

2.3. Objectif thèse, Solutions et Organisation

Comme présenté dans le paragraphe précédent, le problème de l'estimation des données manquantes en raison de la présence de nuages et / ou des ombres représente un domaine de recherche d'un grand intérêt. Dans ce travail de thèse, nous allons nous concentrer sur la reconstruction de données d'image, totalement obscurcies par la présence de nuages ou seulement partiellement masquées par la présence d'ombres, en proposant des solutions méthodologiques nouvelles et innovantes.

Après cette section introductive, la thèse est organisée en six chapitres. Une première partie, contenant trois sections, est consacré au problème de la contamination de l'image par les nuages, puis une deuxième partie, contenant deux sections, traite du problème de la contamination par les ombres.

Dans le chapitre 3, le problème de la reconstruction de zone masquée par la présence de nuages est abordé. En particulier, nous nous concentrons sur la reconstruction basée sur les techniques d'in-painting sur une image d'une seule date. Après une brève introduction de l'algorithme développé par Criminisi et al [36] sur l'inpainting basée région (RBI), trois nouvelles stratégies sont proposées : 1) in-painting sur la base de l'extraction de caractéristiques (FEBI), où des caractéristiques de texture sont exploités; 2) in-painting avec transformation isométrique (ISOI), où nous cherchons à accroître le nombre de candidats de potentiel, et 3) inpainting multirésolution (MRI), où nous avons progressivement injecté les informations multirésolution afin d'obtenir une meilleure reconstruction.

Dans le chapitre 4, une nouvelle stratégie de régression à vecteurs de support (SVR) dans le contexte multitemporel et multispectral des images de télédétection est proposée. En particulier, nous avons amélioré le processus de reconstruction en intégrant à la fois l'information radiométrique et la position spatiale. Pour chaque type d'information utilisé dans la régression, un noyau spécifique est sélectionné et adapté. De façon plus précise, nous évaluons les performances des combinaisons de trois différentes typologies de noyaux (par exemple, linéaire-linéaire, linéaire-

polynomiale, etc.) Par la suite, leur fusion est réalisée par une combinaison linéaire des deux noyaux résultants. La régression par vecteurs supports est alors appliquée pour calculer le fonction de prédiction.

Dans le chapitre 5, trois nouvelles méthodes de reconstruction des données manquantes dues à la présence de nuages sont proposées. Étant donné un ensemble de données multidates, les mesures manquantes sont estimées par application d'un échantillonnage compressé (Compressive Sensing, CS) dans lequel les pixels sans nuages sont exploités. De façon plus détaillée, nous avons d'abord utilisé deux des méthodes le plus courantes de la littérature, à savoir Basis Pursuit (BP) et Orthogonal Matching Pursuit (OMP). Ensuite, nous proposons une solution alternative d'échantillonnage compressé, qui exploite les capacités de recherche des algorithmes génétiques (AG).

Venons-en maintenant à la deuxième partie de la thèse, via le chapitre 6, dans laquelle une nouvelle solution permettant de s'affranchir de la présence des ombres en très haute résolution (THR) est introduite. Nous proposons une chaîne de traitement complète, qui repose sur un traitement d'image avancé et divers outils de reconnaissance de formes. Plus en détails, la détection des zones d'ombre est réalisée par une classification, mise en oeuvre par une méthode à vecteurs supports (support vector machine, SVM), tandis que la reconstruction est basée sur une méthode de régression linéaire, capable d'ajuster les intensités des pixels à l'ombre par rapport à leur correspondants de la même classe en dehors de l'ombre.

Dans le chapitre 7, différents critères sont utilisés pour aider à savoir a priori s'il est possible ou non de reconstruire une zone d'ombre spécifique. En particulier, nous supposons qu'un critère de potentiel de reconstruction idéal ne devrait pas tolérer qu'une zone d'ombre irrécupérable soit considérée comme reconstructible et, en même temps, celui-ci devrait permettre de maximiser la probabilité de détection des zones reconstructible. Plusieurs critères d'évaluation qui travaillent au niveau des pixels et de la texture sont présentés et discutés, afin d'aboutir à la définition d'un indice global basé sur la fusion de deux critères simples qui sont la divergence de Kullback-Leibler des distributions locales et la différence des seconds moments angulaires.

Enfin, dans le dernier chapitre, des considérations finales sur les techniques proposées sont données, avec des questions ouvertes et des travaux de recherche futurs.

2.4. Références

[1] Jian Guo Liu, and Philippa J. Mason, *Essential Image Processing for GIS and Remote Sensing*, Chichester: Wiley-Blackwell, 2009.

[2] Robert A. Schowengerdt, *Remote sensing: Models and methods for image processing*, 3rd ed., Tucson: Academic Press, 2007.

[3] John Robert Schott, *Remote sensing: The image chain approach*, 2nd ed., New York: Oxford University Press, 2007.

[4] R. Hellermann, "Growing demand for VHR satellite imagery for civilian use," in *Proc. International Commercial Remote Sensing Symposium*, Washington, USA, Mar. 2010.

[5] K. Jacobsen, "Very high resolution satellite images – competition to aerial images," in *Proc. Geospatial World Forum*, Hyderabad, India, Jan. 2011.

[6] A. P. Carleer, O. Debeir, and E. Wolff, "Assessment of very high spatial resolution satellite image segmentations," *J. of the American Society for Photogrammetry and Remote Sensing*, vol. 71, pp. 1285–1294, Nov. 2005.

[7] M. Neubert and G. Meinel, "Atmospheric and terrain correction of IKONOS imagery using ATCOR3," in *Proc. ISPRS*, Hannover, Germany, May 2005.

[8] J. R. Wang, P. Racette, M. E. Triesky, E. V. Browell, S. Ismail, and L. A. Chang, "Profiling of atmospheric water vapor with MIR and LASE," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1211–1219, Jun. 2002.

[9] B. G. Vasudevan, B. S. Gohil, and V. K. Agarwal, "Backpropagation neural-network-based retrieval of atmospheric water vapor and cloud liquid water from IRS-P4 MSMR," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 5, pp. 985–990, May 2004.

[10] T. Kim, T. Javzandulam and T-Y. Lee, "Semiautomatic reconstruction of building height and footprints from single satellite images," in *Proc. IGARSS*, Barcelona, Spain, Jul. 2007, pp. 4737–4741.

[11] L. L. Stowe, P. A. Davis, and E. P. McClain, "Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the advanced very high resolution radiometer," *J. Atmospheric and Oceanic Technology*, vol. 16, pp. 656–681, Jun. 1999.

[12] A. V. Di Vittorio, and W. J. Emery, "An automated, dynamic threshold cloud-masking algorithm for daytime AVHRR images over land," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1682–1694, Aug. 2002.

[13] F. Murtagh, D. Barreto, and J. Marcello, "Decision boundaries using Bayes factors: The case of cloud masks," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2952–2958, Dec. 2003.

[14] J. Liu, J. Yang, and T. Fang, "Color property analysis of remote sensing imagery," *Act Photonica Sinica*, vol. 38, no. 2, pp. 441–446, Feb. 2009.

[15] G. Finlayson, and S. Süsstrunk, "Optimization for Hue constant RGB sensors," in *Proc. IS&T/SID*, Portland, OR, USA, Apr. 2002, vol.10, pp. 343–348.

[16] V. Tsai, "A comparative study on shadow compensation of color aerial images in invariant color models," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no.6, pp. 1661–1671, Jun. 2006.

[17] K. L. Chung, Y. R. Lin, and Y. H. Huang, "Efficient shadow detection of color aerial images based on successive thresholding scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 671–682, Feb. 2009.

[18] D. Cai M. Li, Z. Bao, Z. Chen, W. Wei, and H. Zhang, "Study on shadow detection method on high resolution remote sensing image based on HIS space transformation and NDVI index," in *Proc. International Conference on Geoinformatics*, Beijing, China, Jun. 2010, pp.1–4.

[19] H. Ma, Q. Qin, and X. Shen, "Shadow segmentation and compensation in high resolution satellite images," in *Proc. IGARSS*, Boston, MA, USA, Jul. 2008, vol. 2, pp. 1036–1039.

[20] H.Y. Yu, J.G. Sun, L.N. Liu, Y.H. Wang, and Y.D. Wang, "MSER based shadow detection in high resolution remote sensing image," in *Proc. ICMLC*, Qingdao, China, Jul. 2010, pp. 780–783.

[21] S. Wang, and Y. Wang, "Shadow detection and compensation in high resolution satellite images based on retinex," in *Proc. ICIG*, Xi'an, China, Sep. 2009, pp. 209–212.

[22] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signals Process. Mag.*, vol. 14, no. 2, pp. 24–41, Mar. 1997.

[23] S. E. Reichenbach, D. E. Koehler, and D. W. Strelow, "Restoration and reconstruction of AVHRR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 997–1007, Jul. 1995.

[24] V. P. Bakalov and M. Y. Yerokhin, "Removal of uncontrollable phase distortions in synthetic aperture radar signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1298–1302, May 2000.

[25] S. E. Reichenbach and J. Li, "Restoration and reconstruction from overlapping images for multi-image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 4, pp. 769–780, Apr. 2001.

[26] Z. Wu and C. Liu, "An image reconstruction method using GPR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 327–334, Jan. 1999.

[27] S. C. Liew, M. Li, and L. K. Kwoh, "Automated production of cloud-free and cloud-shadow image mosaics from cloudy satellite imagery," in *Proc. ISPRS*, Istanbul, Turkey, Jul. 2004, pp. 523–530.

[28] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens*., vol. 44, no. 2, pp. 442–455, Feb. 2006.

[29] B.-C. Gao, P. Yang, W. Han, R.-R. Li, and W. J. Wiscombe, "An algorithm using visible and 1.38-μm channels to retrieve cirrus cloud reflectances from aircraft and satellite data," *IEEE Trans. Geosci. Remote Sens*., vol. 40, no. 8, pp. 1659–1668, Nov. 2002.

[30] A. Maalouf, P. Carré, B. Augereau and C. F. Maloigne, "A bandelet-based inpainting technique for clouds removal from remotely sensed images," *IEEE Trans. Geosci. Remote Sens*., vol. 47, no. 7, pp. 2363–2371, Jul. 2009.

[31] P. Sarabandi, F. Yamazaki, M. Matsuoka, and A. Kiremidjian, "Shadow detection and radiometric restoration in satellite high resolution images," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 6, pp. 3744–3747.

[32] A. Massalabi, H. Dong-Chen, G.B. Benie and E. Beaudry, "Detecting information under and from shadow in panchromatic IKONOS images of the city of Sherbrooke," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 3, pp. 2000–2004.

[33] J. Su, X. Lin, and D. Liu, "An automatic shadow detection and compensation method for remote sensed color images," in *Proc. ICSP*, Beijing, China, Nov. 2006, vol. 2.

[34] T. Nakajima, G. Tao, and Y. Yasuoka, "Simulated recovery of information in shadow areas on IKONOS image by combining ALS data," in *Proc. ACRS*, Kathmandu, Nepal, 2002.

[35] F. Yamazaki, W. Liu, and M. Takasaki, "Characteristic of shadow and removal of its effects for remote sensing imagery," in *Proc. IGARSS*, Cape Town, South Africa, Jul. 2009, vol. 4, pp.426–429.

[36] A. Crimisi, P. Perez and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. on Image Process*., vol. 13, no. 9, pp. 1–14, Sep. 2004.

3. Inpainting Strategies for Reconstruction of Missing Data in VHR Images

Abstract *– Missing data in very high spatial resolution (VHR) optical imagery take origin mainly from the acquisition conditions. Their accurate reconstruction represents a great methodological challenge because of the complexity and the ill-posed nature of the problem. In this chapter, we present three different solutions, all based on the inpainting approach, which consists in reconstructing the missing regions in a given image by propagating the spectro-geometrical information retrieved from the remaining parts of the image. They rely on the idea to enrich the patch search process by including local image properties or by isometric transformations, or to reformulate it under a multiresolution processing scheme, respectively. Thorough experiments conducted on two different VHR images are reported and discussed.*

The work presented in this chapter has been published in the *IEEE Geosci. and Remote Sens. Lett.*, vol. 8, no. 6, pp. 914–918, September 2011; Co-authors: F. Melgani and G. Mercier.

3.1. Introduction

In the last years, different works have been presented to cope with the problem of the reconstruction of missing data due mainly to cloud covers in passive imagery. Most of these solutions are intended for low or medium spatial resolution images and are based on the assumption that the temporal signature of a given pixel is contaminated by residual effects caused by imperfect sensing of the target or by spatially autocorrelated noise due to atmospheric attenuation. They make use of principles such as the data substitution principle [1] or temporal prediction strategies [2]-[3]. Their common denominator, which consists in the need for a sequence of (high temporal resolution) images, can however sometimes be viewed as a limitation.

As an alternative, one can resort to the filling-in approach such as inpainting techniques, which aim at filling holes in digital images by propagating surrounding structures. An early paper dealing with digital inpainting introduces basic techniques inspired from tricks of professional restorers of old paints [4]. Another work proposes to synthesize a complete, visually plausible and coherent image by applying an improved fast fragment-based image completion technique [5]. Other interesting papers based on the wavelet transformation can also be found in the literature [6]-[7], where the authors apply inpainting at different wavelet decomposition levels and obtain the reconstructed image by inverse transformation. More recently, in [1], an inpainting method allows removing clouds by means of the bandelet transform (a special case of the wavelet transform) and the multiscale geometrical grouping. Its promising experimental results encourage the idea of exploiting inpainting techniques to reconstruct missing data in remote sensing imagery.

In this chapter, we present various inpainting alternatives for the reconstruction of missing areas in images specifically acquired by very high spatial resolution (VHR) sensors whose images are quickly growing in popularity among the remote sensing community.

3.2. Problem Formulation

Let us consider a VHR image $I[x, y]$ with dimension of $N \times M$ pixels and partly contaminated. We assume the missing part has been beforehand localized by an opportune technique. Image I can therefore be divided in two parts: a *target region* Ω, characterized by missing data, and a *source region* Φ, from where the most similar patches to reconstruct Ω are extracted:

$$I = \Omega \cup \Phi. \tag{3.1}$$

The boundary between target and source regions is called *fill front* and is indicated with $\delta\Omega$. The scope of any inpainting technique is, starting from the fill front, to fill in Ω using information contained in Φ (see Figure 3.1(a)). One of the inpainting modalities relies on patches instead of single pixels since the pixel-based inpainting approach is not able to correctly synthesize original image textural properties [9]-[10]. A patch is a small square region $\Psi_{\mathrm{p}}[x, y]$ centered in $\mathrm{p}[x, y]$ and containing $W \times W$ pixels (see Figure 3.1(b)). The problem is thus to find the most similar patch $\Psi_{\mathrm{q}}[x, y] \in \Phi$ (see Figure 3.1(c)). Once found, it is pasted to substitute the patch $\Psi_{\mathrm{p}}[x, y] \in \Omega$ (see Figure 3.1(d)). This filling approach is also called region-based, where the region in this case is the

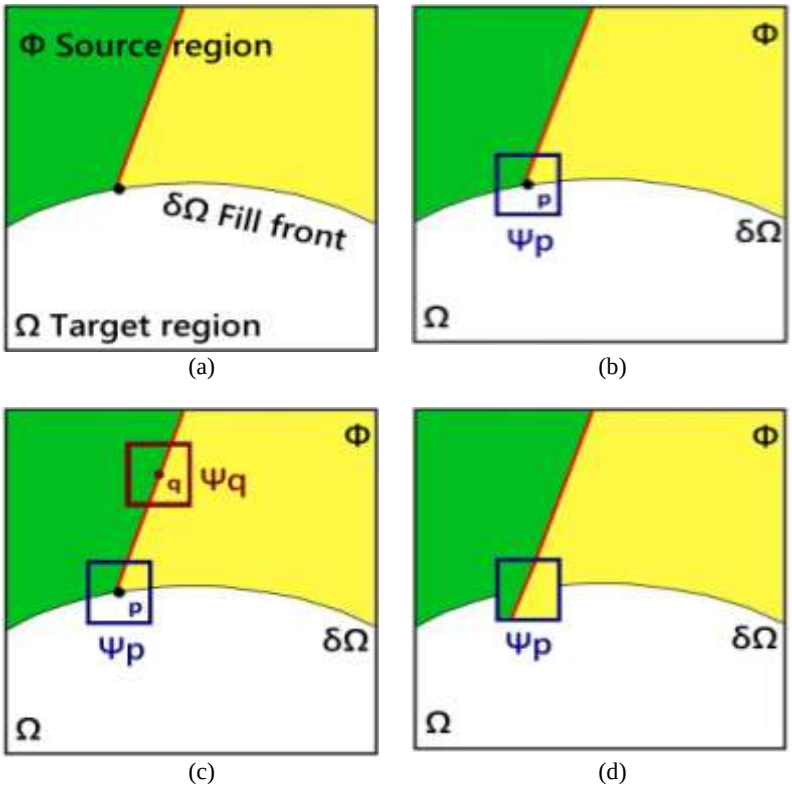


Figure 3.1. Illustration of the fill-in propagation by region-based synthesis. (a) The original image is divided in source region (composed of green and yellow areas) and target region (white area). (b) A patch $\Psi_p[x, y]$ is centered in $p[x, y]$ on the fill front $\delta\Omega$. (c) The most similar patch $\Psi_q[x, y] \in \Phi$. (d) Inpainting result for $\Psi_p[x, y]$.

patch Ψ. This inpainting procedure is recursively iterated until all pixels from the target region Ω are reconstructed. In the following, more details are provided.

3.2.1. Region-based inpainting (RBI)

This chapter proposes different algorithms taking origin from a previous work from Crimisi *et al*. [11], which is considered as one of the state-of-the-art inpainting algorithms. Their original work is region-based and assigns to each pixel $p[x, y] \in \delta\Omega$ an inpainting priority value $P(p)$ defined by:

$$P(p) = C(p)D(p), \tag{3.2}$$

where $C(p)$ and $D(p)$ are *confidence* and *data terms*, respectively, given by:

$$C(p) = \frac{\sum_{q \in \Psi_p \cap (I-\Omega)} C(q)}{|\Psi_p|} \text{ and } \quad D(p) = \frac{|\nabla I_p^{\perp} \cdot \mathbf{n}_p|}{2^n} \tag{3.3}$$

$|\Psi_p|$ simply corresponds to the area of the patch Ψ_p, $\mathbf{n}_p$ is the unit vector orthogonal to the border $\delta\Omega$ at the pixel p, $\nabla I_p^{\perp}$ denotes the orthogonal vector of the local gradient ∇I_p and 2^n is a normalization factor, where $n = 8$ for a typical grey-level image (see Figure 3.2). The parameter $C(p)$ represents a confidence value, which expresses the amount of reliable information surrounding the pixel p. During the initialization, it can be evaluated as:

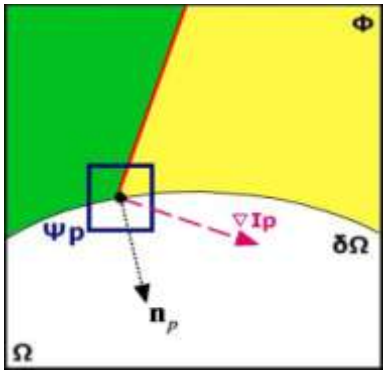


Figure 3.2. Inpainting geometry. Patch $\Psi_p[x,y]$ centered in $p[x,y]$ on the fill front $\delta\Omega$. n_p: unit vector orthogonal to the border $\delta\Omega$ in the pixel p. ∇I_p: vector orthogonal to the isophote at point p.

$$\begin{cases} C(p) = 0 & \forall p \in \Omega \\ C(p) = 1 & \forall p \in \Phi = I - \Omega \end{cases} \tag{3.4}$$

The parameter $D(p)$ considers instead the continuation of strong edges. For the detailed definitions and the mathematical derivations, we refer the reader to [11].

The region-based inpainting (RBI) algorithm can be summarized as follows:

1. Initialize the confidence term $C(p) = 0 \ \forall p \in \Omega$ and $C(p) = 1 \ \forall p \in \Phi = I - \Omega$ (note that $\delta\Omega^0 \in \Omega$);
2. Repeat until $\Omega^t = \emptyset$:
   - *i.* Compute priority $P(p) \ \forall p \in \delta\Omega^t$
   - *ii.* Find the patch $\Psi\hat{p}$ such that $\hat{p} = \max_{p \in \delta\Omega^t} P(p)$
   - *iii.* Find $\Psi\hat{q} \in \Phi$ so that $\Psi\hat{q} = \min_{\Psi_q \in \Phi} d(\Psi\hat{p}, \Psi q)$
   - *iv.* Copy $\Psi\hat{q}$ into $\Psi\hat{p}$, $\forall p \in (\Psi\hat{p} \cap \Omega)$
   - *v.* Update $C(p) = C(\hat{p}), \ \forall p \in (\Psi\hat{p} \cap \Omega)$ and the border $\delta\Omega^t$
   - *vi.* Increment t

In the original algorithm, the distance between two patches, which has to be minimized in *Step 2.iii*, is evaluated as follows:

$$d\left(\Psi_{\hat{p}}[x,y], \Psi_q[u,v]\right) = \sum_{(i,j):(x-i,y-j)\in\Phi} \left(\Psi_{\hat{p}}[x-i, y-j] - \Psi_q[u-i, v-j]\right)^2 \tag{3.5}$$

In other words, it computes the well-known sum square error (SSE) between the two patches $\Psi\hat{p}$ and Ψq. One of the limitations of the region-based inpainting (RBI) algorithm is its sensitivity to the size of Ω. In the following, we propose different enhancements in an attempt to improve the effectiveness of the reconstruction process in terms of both the accuracy as well as the robustness to the problem of target region size.

3.3. Proposed Inpainting Strategies

The proposed strategies move in three different directions. The first one consists to enrich the feature space so that to refine the search for the most similar patch in the source region. This will be done by adding textural features extracted from the original image. Through isometric transformations, the second direction aims at increasing the amount of patch candidates and thus populating further the search space. The last direction faces the inpainting problem within a multiresolution processing scheme.

3.3.1. Feature extraction-based inpainting (FEBI)

As mentioned before, the first strategy intends to facilitate the search for the most similar patch Ψq by considering not just original image channels but also image features, which capture local image properties potentially useful for the best patch identification. Operationally speaking, this means applying RBI with a similarity function ($d(\Psi\hat{p}, \Psi q)$) that integrates the extracted image features.

In this work, two kinds of image features are considered. The first expresses the local variance of the image (*Stdv*), while the second kind is a space-frequency representation of the original image by means of the symlet transform (*Sym*), which is a particular case of the wavelet transform [12]. In both cases, the similarity function (3.5) to minimize can be rewritten as:

$$d\left(\Psi_{\hat{p}}[x,y], \Psi_q[u,v]\right) = \quad (3.6)$$
$$= \sum_{(i,j):(x-i,y-j)\in\Phi} \left(\Psi_{\hat{p}}[x-i,y-j] - \Psi_q[u-i,v-j]\right)^2 + \alpha_i\left(F_{\hat{p}}[x-i,y-j] - F_q[u-i,v-j]\right)^2$$

where $F_i[x,y]$ corresponds to the extracted feature (i.e., *Stdv* or *Sym*) and α_i is a weighting parameter which controls the influence of the extracted features on the search process.

3.3.2. Inpainting with isometric transformation (IsoI)

One of the potential problems of RBI is that the patch search is directionally constrained by the spatial image structure. By contrast, patches could be searched for in various directions for a better mining of the inpainting possibilities conveyed in the image. A second strategy we explore consists therefore to implement a patch search process by isometric geometrical transformations [13]. This means that the source patch Ψ_q will be transformed (rotated and/or flipped) to match better the target patch $\Psi_{\hat{p}}$. In this work, a subset of seven common transformations is selected (see Figure 3.3).

In IsoI, the distance measure between two patches becomes:

$$d\left(\Psi_{\hat{p}}[x,y], iso_i\left(\Psi_q[x,y]\right)\right) = \sum_{(i,j):(x-i,y-j)\in\Phi}\left(\Psi_{\hat{p}}[x-i,y-j] - iso_i\left(\Psi_q[x-i,y-j]\right)\right)^2, \quad (3.7)$$

where $iso_i\left(\Psi_q\right)$ stands for an isometric transformation of Ψ_q.

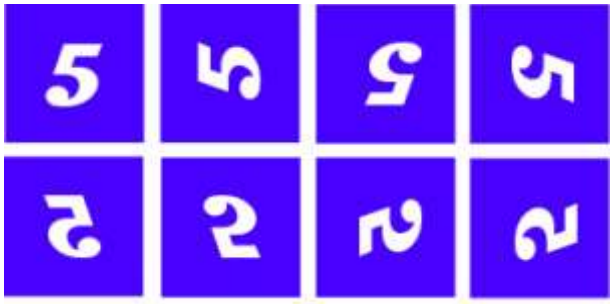


Figure 3.3. Illustration of the 7 isometric transformations: +90°, +180°, +270°, horizontal flip, vertical flip, horizontal flip +90° and horizontal flip +270°.

3.3.3. Multiresolution inpainting (MRI)

The third and last strategy is based on a completely different idea motivated by the success it obtained in different image processing and analysis application fields [14]-[15]. It consists in a multiresolution processing of the image so that to reconstruct missing data with a progressively increasing spatial accuracy. In more detail, the algorithm starts by applying RBI with a patch of large window dimension W to fill in the considered image hole (Figure 3.4). In the next iteration, the size of the patch is decreased and RBI is applied again with a first difference that this time missing data have been substituted by an estimate obtained in the previous iteration. In order to maintain a memory between successive resolution levels, another difference is that instead of performing a simple pasting of the new patch, the reconstruction will be based on a weighted sum of the new patch and the one found at previous resolution. This process is repeated up to reach the predefined smallest value of W.

MRI can be resumed as follows:

Step 1: Fix initial and final values of window size W_1 and W_F, respectively. Set $i = 1$ and $I_i = I$.

Step 2: Apply RBI with W_i to I_i. Compute Ψ_{p_n} as follows:

$\Psi_{p_n} = \Psi_{q_n}$, if $i = 1$

$\Psi_{p_n} = \alpha\Psi_{q_n} + (1-\alpha)\Psi_{p_{n-1}}$, otherwise

Set $i = i + 1$. Store output image in I_i. Set $W_i = W_{i-1}/2$.

Step 3: If $W_i \geq W_F$, go to *Step 2*. Otherwise, stop.

α is a weight factor which governs the "hysteresis" effect resulting from the progressive change of resolution. Under the MRI strategy, it comes out that:

$$\Psi_{p_n} = \alpha\Psi_{q_n} + \alpha(1-\alpha)\Psi_{q_{n-1}} + \cdots + \alpha(1-\alpha)^{n-2}\Psi_{q_2} + (1-\alpha)^{n-1}\Psi_{q_1} \quad (3.8)$$

In other terms, all resolution levels contribute with different degrees to the final reconstruction of the missing data region. Note that MRI is a general strategy which could be implemented as well jointly with the two previous inpainting strategies.

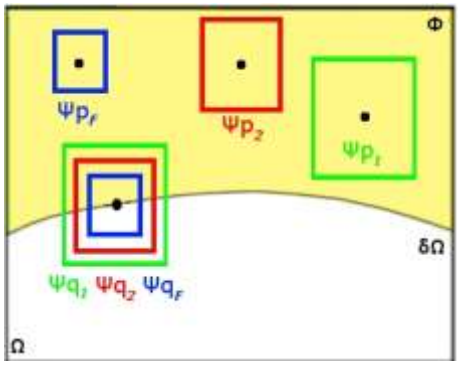


Figure 3.4. In MRI, the window dimension of the patches Ψ_{p_i} and Ψ_{q_i} decreases with iterations.

3.4. Experimental Results

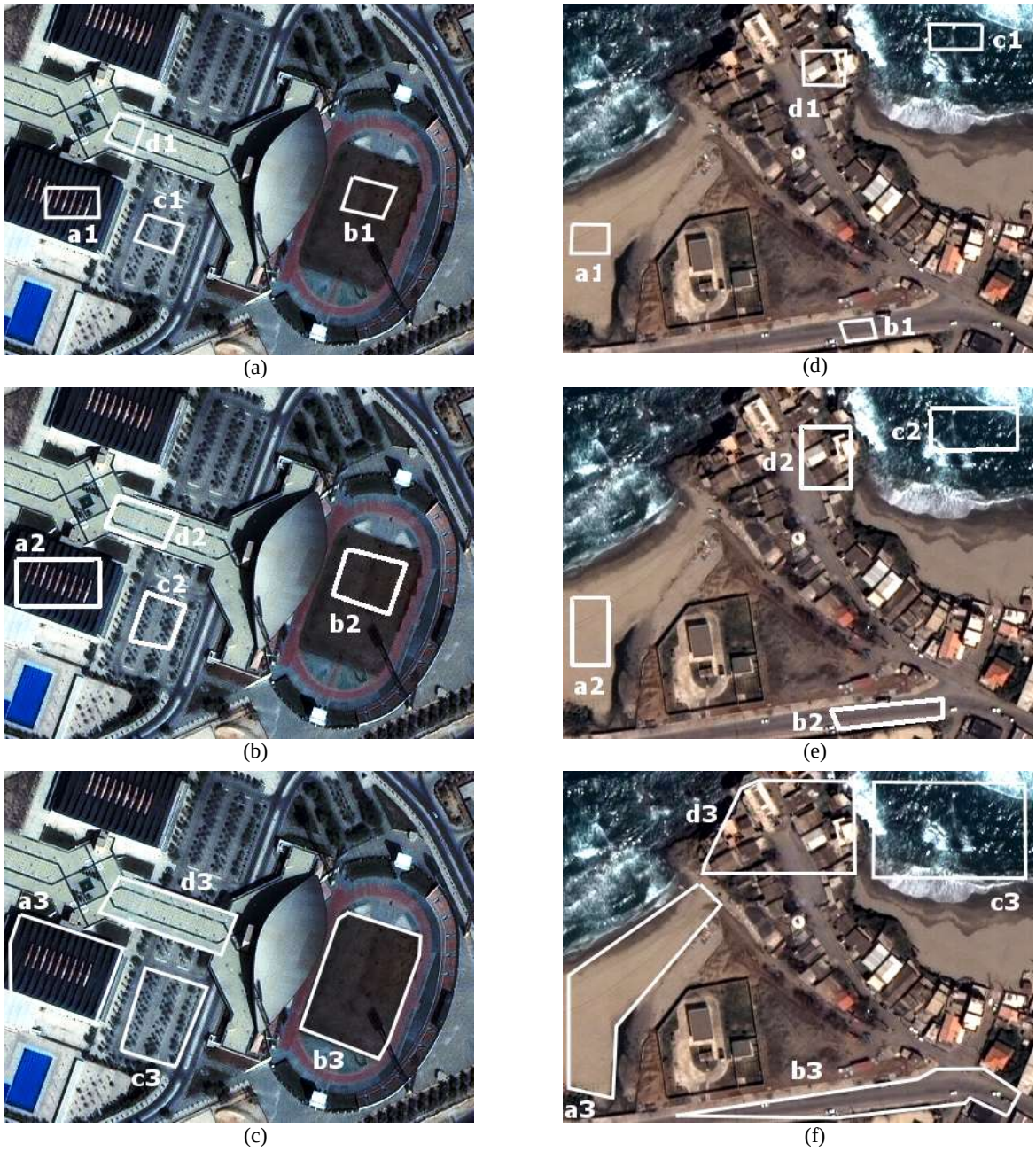
The various simulations we conducted and presented in this section aim at testing the above described inpainting strategies under different contamination conditions, including the typology of the missing area and its size.

3.4.1. Dataset description and setups

Two different cloud-free scenes are considered. The first one refers to an IKONOS image (composed of the RGB spectral bands), representing a part of the city of Riyadh (Saudi Arabia) and acquired on the 28th of April 2008. The second is a QuickBird image (4 spectral bands), acquired on the 23th of May 2003 over a part of the coastal region of Boumerdes (Algeria).

For our experiments, the two scenes have been cropped into 300×400 pixel-size images containing different land cover typologies such as grassy, sandy and rocky areas, and roofs and parking lots. As shown in Figure 3.5, for each image and for each land cover type, a mask of varying size is defined (*1*: small, *2*: medium and *3*: large). The resulting masks make it possible to quantify the reconstruction accuracy by comparing the reconstruction result with the true pixel values. As an accuracy measure, we adopt the well-known peak signal-to-noise ratio (PSNR) measure [16].

The setting of the free parameters α_i of the different strategies is done by empirical estimation, i.e., by maximizing the PSNR over a predefined area of the source region Φ. From a pattern recognition viewpoint, such an area plays the role of validation area useful to tune algorithm parameters.



(a) (d) (b) (e) (c) (f)

Figure 3.5. Riyadh (left) and Boumerdes (right) images with masks of increasing dimension (from top to bottom). Note that Riyadh image has a spatial resolution of 4 meters and the one of Boumerdes is 2.4 meters.

3.4.2. Experimental results

3.4.2.1. Contamination of different land covers

Table 3.I reports the results of the various inpainting strategies applied over different obscured land covers by considering masks of medium size (masks *a2*, *b2*, *c2*, and *d2*). Each mask is handled separately, that is by assuming that it refers to the only contaminated part of the image. In addition, experiments are also performed by supposing multiple land cover contamination, namely different areas of the image are missing. In this case, the corresponding mask is *e2*, which refers to the union of the four masks *a2*, *b2*, *c2*, and *d2*.

The explored strategies generally outperform the reference strategy RBI in terms of PSNR, except for IsoI, which does better only in few cases. Note that MRI works generally better compared with the other methods. Moreover, one can observe that enriching the patch search with additional extracted features can help. Conversely, it appears that the isometric transformations can mislead the inpainting process by pasting patches not necessarily compatible with the geometric configuration of the image. Figures 3.6(b) and 3.6(e) show the reconstruction results obtained by the best strategy, i.e., MRI, for the Riyadh and Boumerdes images contaminated with the mask *e2* (last column of Table 3.I).

TABLE 3.I
RESULTS IN PSNR OBTAINED BY THE DIFFERENT INPAINTING STRATEGIES FOR DIFFERENT LAND COVERS.
(a) RIYADH IMAGE AND (b) BOUMERDES IMAGE.

(a)

Method	***Mask a2***	***Mask b2***	***Mask c2***	***Mask d2***	***Mask e2***
RBI	16.68	29.13	12.33	12.17	15.20
FEBI-Stdv	17.26	29.13	12.78	12.17	15.49
FEBI-Sym	18.47	29.01	12.09	12.65	15.75
IsoI	16.35	23.40	13.71	9.42	14.18
MRI	**19.91**	**29.73**	**16.81**	**14.47**	**18.41**

(b)

Method	***Mask a2***	***Mask b2***	***Mask c2***	***Mask d2***	***Mask e2***
RBI	36.86	30.57	20.48	20.81	22.90
FEBI-Stdv	41.23	26.91	22.88	21.07	23.70
FEBI-Sym	35.90	**30.66**	23.06	20.48	**23.80**
IsoI	29.93	23.40	22.10	19.72	22.09
MRI	**44.35**	28.34	**23.09**	**21.67**	23.35

3.4.2.2. Contamination of different sizes

Another important test for the inpainting strategies consists in assessing their performance by varying the amount of missing data. Here we will report the results just for two kinds of land covers, i.e., homogeneous and heterogeneous land covers.

For such purpose, the mask *b* covering grass areas in the Riyadh image and the mask *a* representing sandy areas in the Boumerdes image have been selected. For each mask, we run the strategies by considering this time small and large mask sizes. The obtained quantitative results are listed in Table 3.II, while qualitative results for MRI are shown in Figure 3.6. For small sizes, all methods perform well. As it could be expected due to intrinsic propagation errors incurred by the inpainting approach, for large sizes, the reconstruction accuracy drops drastically. It is noteworthy that the proposed strategies exhibit a relatively higher robustness to this issue.

TABLE 3.II
RESULTS IN PSNR OBTAINED BY THE DIFFERENT INPAINTING STRATEGIES BY VARYING THE MISSING AREA SIZE IN HOMOGENEOUS LAND COVERS. (a) RIYADH IMAGE AND (b) BOUMERDES IMAGE.

(a)

Method	*Mask b1*	*Mask b2*	*Mask b3*
RBI	29.98	29.13	10.18
FEBI -Stdv	29.98	29.13	11.13
FEBI -Sym	29.79	29.01	12.12
IsoI	28.52	23.40	12.30
MRI	**31.17**	**29.73**	**13.97**

(b)

Method	*Mask a1*	*Mask a2*	*Mask a3*
RBI	42.76	36.86	24.95
FEBI -Stdv	42.60	41.23	26.47
FEBI -Sym	43.93	35.90	25.46
IsoI	43.05	29.93	25.26
MRI	**44.56**	**44.35**	**26.98**

For the test on heterogeneous areas, we adopted the mask *a*, which obscures roofs in the Riyadh image, and the mask *d*, which covers an urban area in the Boumerdes image. Results are reported in Table 3.III. In general, similar conclusions as in the previous experiments can be drawn. For the Boumerdes image, an increase of PSNR was possible since the mask *d3* covers also a portion of homogenous area compared to *d1*, which is completely superimposed over roofs. The accuracy decrease from the smallest mask to the largest one is smaller compared to the previous experiments since the PSNR starts with relatively low values due to the higher complexity of the obscured areas.

Finally, in Table 3.IV, a comparison of the different strategies in terms of how many times they return better, worse or similar results compared with RBI over all the thirty performed experiments is reported. Most of the algorithms return better results compared with RBI, except IsoI. From this table, it clearly appears that the best method is MRI.

3.5. Conclusion

This chapter deals with the complex and important problem to reconstruct VHR images affected by the presence of missing areas. In particular, different inpainting strategies are proposed, which allow to reconstruct the contaminated regions by propagating the spectro-geometrical information retrieved from outside the missing area.

All strategies exhibit the advantage of being independent from the sensor type and from its spatial, temporal and spectral properties. Moreover, they are completely unsupervised. Their major drawback is their sensitivity to the size of the missing area, a drawback inherited by their inpainting-based nature. Indeed, it is important that the source region contains information which is missing in the target region. The risk that it does not include such information, increases as the size

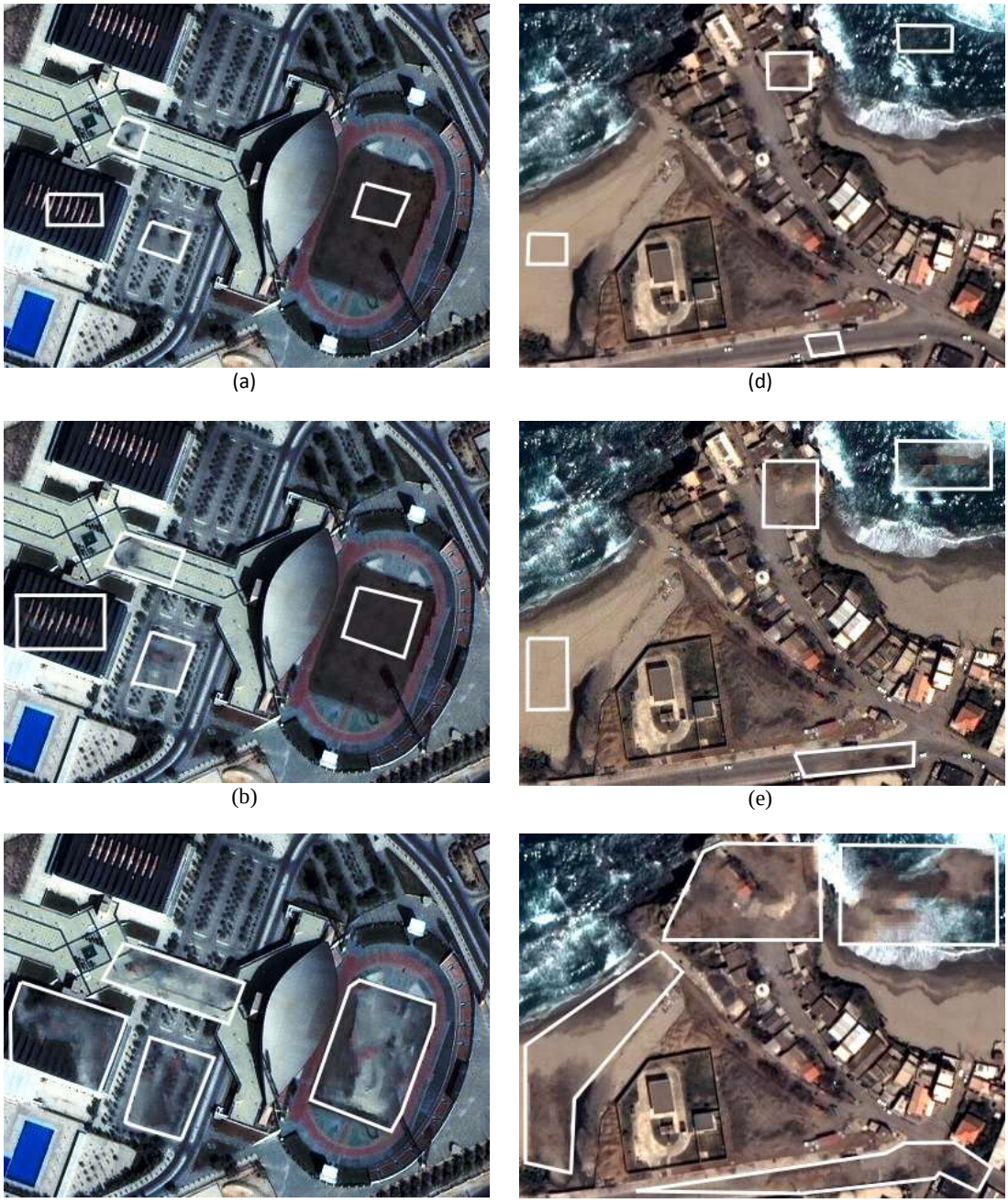


Figure 3.6. Reconstruction results obtained by the MRI method for the masks illustrated in Figure 3.5. Riyadh image (a)-(c) and Boumerdes image (d)-(e).

of the target region augments. From the results, one can deduce that inpainting can be improved by enriching the patch search with local image features and that isometric transformations can mislead the search process. The best inpainting solution is MRI because of its progressive and multiresolution approach to the reconstruction problem.

Finally, in order to improve the accuracy of the reconstruction process, the integration of the temporal dimension in the inpainting approach deserves to be investigated in a future research study.

TABLE 3.III
RESULTS IN PSNR OBTAINED BY THE DIFFERENT INPAINTING STRATEGIES BY VARYING THE MISSING AREA SIZE IN HETEROGENEOUS LAND COVERS. (a) RIYADH IMAGE AND (b) BOUMERDES IMAGE.

(a)

Method	*Mask a1*	*Mask a2*	*Mask a3*
RBI	19.60	16.68	11.73
FEBI -Stdv	19.60	17.26	12.29
FEBI -Sym	19.70	18.47	12.17
IsoI	19.60	16.35	10.42
MRI	**20.13**	**19.91**	**13.33**

(b)

Method	*Mask d1*	*Mask d2*	*Mask d3*
RBI	19.99	20.81	21.88
FEBI -Stdv	19.82	21.07	21.77
FEBI -Sym	**20.05**	20.48	21.49
IsoI	19.85	19.72	21.83
MRI	19.39	**21.67**	**22.39**

TABLE 3.IV
GLOBAL COMPARISON BETWEEN PROPOSED STRATEGIES AND RBI.

Method	Better	Worse	No change
FEBI-Stdv	12/30	11/30	7/30
FEBI-Sym	18/30	11/30	1/30
IsoI	9/30	20/30	1/30
MRI	**27/30**	**3/30**	0/30

3.6. Acknowledgment

The authors would like to thank Dr. Y. Bazi (King Saud University, Saudi Arabia) for providing the Riyadh image.

3.7. References

[1] S. C. Liew, M. Li, and L. K. Kwoh, "Automated production of cloud-free and cloud-shadow image mosaics from cloudy satellite imagery," in *Proc. XXth ISPRS*, Istanbul, Turkey, Jul. 2004, pp. 523–530.

[2] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 442–455, Feb. 2006.

[3] S. Benabdelkader and F. Melgani, "Contextual spatiospectral postreconstruction of cloud-contaminated images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 204–208, Apr. 2008.

[4] M. Bertalmio and G. Sapiro, "Image inpainting," in *Proc. SIGGRAPH*, Los Angeles, CA, USA, 2000, pp. 417–424.

[5] F. Chen, Z. Zhao, L. Peng, and D. Yan, "Clouds and cloud shadows removal from high-resolution remote sensing images," in *Proc. IGARSS*, Seoul, South Korea, Jul. 2005, vol. 6, pp. 4256–4259.

[6] D. Cho and T.D. Bui, "Image inpainting using wavelet-based inter- and intra-scale dependency," in *Proc. ICPR*, Tampa, FL, USA, Dec. 2008, pp. 1–4.

[7] U. A. Ignácio and C. R. Jung, "Block-based image inpainting in the wavelet domain," *Visual Comput.*, vol. 23, no. 9, pp. 733–741, Jun. 2007.

[8] A. Maalouf, P. Carré, B. Augereau and C. F. Maloigne, "A bandelet-based inpainting technique for clouds removal from remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2363–2371, Jul. 2009

[9] A. Hirani and T. Totsuka, "Combining frequency and spatial domain information for fast interactive image noise removal," in *Proc. SIGGRAPH*, New Orleans, LA, USA, Aug. 1996, pp. 269–276.

[10] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. ICCV*, Corfu, Greece, Sep. 1999, pp. 1033–1038.

[11] A. Crimisi, P. Perez and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. on Image Process.*, vol. 13, no. 9, pp. 1–14, Sep. 2004.

[12] R. Singh, R.E. Vasquez and R. Singh, "Comparison of Daubechies, Coiflet, and Symlet for edge detection," in *Proc. SPIE Visual Information Processing VI*, Orlando, FL, USA, Apr. 1997, vol. 3074, pp. 151–159.

[13] Y. Fisher, *Fractal image compression: Theory and application*. New York: Springer-Verlag, 1995.

[14] C. Wemmert, A. Puissant, G. Forestier, P. Gancarski, "Multiresolution remote sensing image clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 533–537, Jul. 2009.

[15] Y. Bazi, F. Melgani, and H. Al-Sharari, "Unsupervised change detection in multispectral remotely sensed imagery with level set methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 8, pp. 3178–3187, Aug. 2010.

[16] A.K. Jain, *Fundamentals of digital image processing*. New York: Prentice Hall, 1988.

4. Support Vector Regression with Kernel Combination for Missing Data Reconstruction

Abstract *– Over the past few years, the reconstruction of missing data due to the presence of clouds received an important attention. Applying region-based inpainting strategies or conventional regression methods, such as support vector machine (SVM) regression, may not be the optimal way. In this chapter, we propose new combinations of kernel functions with which we obtain a better reconstruction. In particular in the regression, we add to the radiometric information, the position information of the pixels in the image. For each kind of information adopted in the regression, a specific kernel is selected and adapted. Adopting this new kernel combination in a support vector regression (SVR), comes out that only few support vectors (SVs) are needed to reconstruct a missing area. This means that we also perform a compression in the number of values needed for a good reconstruction. We illustrate the proposed approaches through some simulations on FORMOSAT-2 multitemporal images.*

The work presented in this chapter has been accepted to be published in the *IEEE Geosci. and Remote Sens. Lett.*; Co-authors: F. Melgani and G. Mercier.

4.1. Introduction

The presence of clouds in remote sensing optical images may produce missing data. In general, if the application does not need to study them, as in the present chapter, clouds represent only an unwanted noise, which distorts the spectral response of land covers. In the last years, different works have been presented to cope with this problem. Usually the solutions are intended to detect and to remove cloud presences, for low or medium spatial resolutions and may require temporal information [1]-[3]. Note that the focus of this chapter is on the reconstruction of the missing areas; their detection is thus not considered here. One of the first techniques which deals with this problem produces a cloud-free image mosaic from several cloudy optical satellite images acquired from the same area [4]. Here the goal is to compose a reasonably cloud-free composite scene by merging together different parts of images. In a similar way, in [5], authors use a regression trees strategy and a histogram matching method to obtain a more plausible mosaic image. Other techniques to reconstruct missing areas are based on prediction techniques. A first work was presented in [6], where a least-squares linear prediction with escalator structure is implemented. In this case, the algorithm predicts each missing pixel from its temporal behavior. Furthermore, unsupervised contextual prediction has been adopted, as in [7], where the local spectro-temporal relationship is used to predict the missing data through a nonlinear regressor. In particular, the author adopts a support vector machine regressor (SVR). To better exploit available information, authors of [8] make use of the spatial and spectral correlations. The literature reports also another technique which exploits SAR information to remove the cloud presence in optical imagery [9]. In [10], an inpainting method is introduced for removing clouds by means of the bandelet transform (a special case of the wavelet transform) and a multiscale geometrical grouping. In [11], several region-based inpainting strategies are proposed to reconstruct missing regions by propagating spectral and geometric information from the remaining parts of the image. The underlying idea is to enrich the region (patch) search process by including local image properties, by isometric transformations, or to reformulate it under a multiresolution processing scheme.

In this chapter, we propose to improve the reconstruction process by integrating both radiometric and position information. For each kind of information, a specific kernel is selected and adapted. Subsequently, their fusion is performed by a linear combination of the two resulting kernels. Support vector regression is applied to derive the prediction function [12]. Simulations conducted on temporal images acquired with the optical high resolution FORMOSAT-2 satellite are reported and discussed.

The remainder of this chapter is organized as follows. In the next section, we will formulate the reconstruction problem. In Section 4.3, we briefly introduce the SVR theory adopting classical kernels and we propose new kernel combinations. The efficiency of the proposed approach is illustrated in Section 4.4 and Section 4.5 draws the conclusions.

4.2. Problem Formulation

Let us consider multitemporal *B*-bands data acquired and registered over the same geographical area by an optical sensor at two different dates, $I_b^{(1)}$ and $I_b^{(2)}$, with $b \in \{1,2,\dots,B\}$ the set of spectral bands. We assume that the acquisitions of the images are temporally close to each other and they are characterized by similar spatial structures. As mentioned above, we remark that the detection step, useful to find the position of the clouds, is not treated in this chapter. We make the hypothesis that image $I^{(2)}$ is obscured by the presence of a cloud. This cloudy area in image $I^{(2)}$ is viewed as a target region Ω and the remaining part as the source region Θ, s.t., $I^{(2)} = \Omega^{(2)} \cup \Theta^{(2)}$ (following classical notation in the inpainting literature). Note that image $I^{(1)}$ is supposed cloud-free. If the case it is not, the proposed technique could be applied in a similar way from $I^{(2)}$ to $I^{(1)}$. The only restriction is that the clouds need not to be at the same positions in the image set.

The aim of this work is to generate a new image $Y^{(2)}$ such that for a specific band *b* and for each pixel in the coordinates (k,l):

$$Y_b^{(2)}(k,l) = \begin{cases} I_b^{(2)}, & if \quad (k,l) \in \Theta^{(2)} \\ f\left((k,l), I^{(1)}\right), & otherwise \end{cases} \tag{4.1}$$

where $f(\cdot)$ represents a prediction function which takes into account the position coordinates (k,l) of a pixel and radiometric information of $I^{(1)}$. The fact to take the location information (k,l) in addition to the radiometric information takes its inspiration from the idea to promote the use of support vectors (SVs) in the same location as the pixel to be inferred. It helps in considering the same kind of landscape in the missing data reconstruction process. Figure 4.1, illustrates the reconstruction process: 1) training and 2) prediction steps.

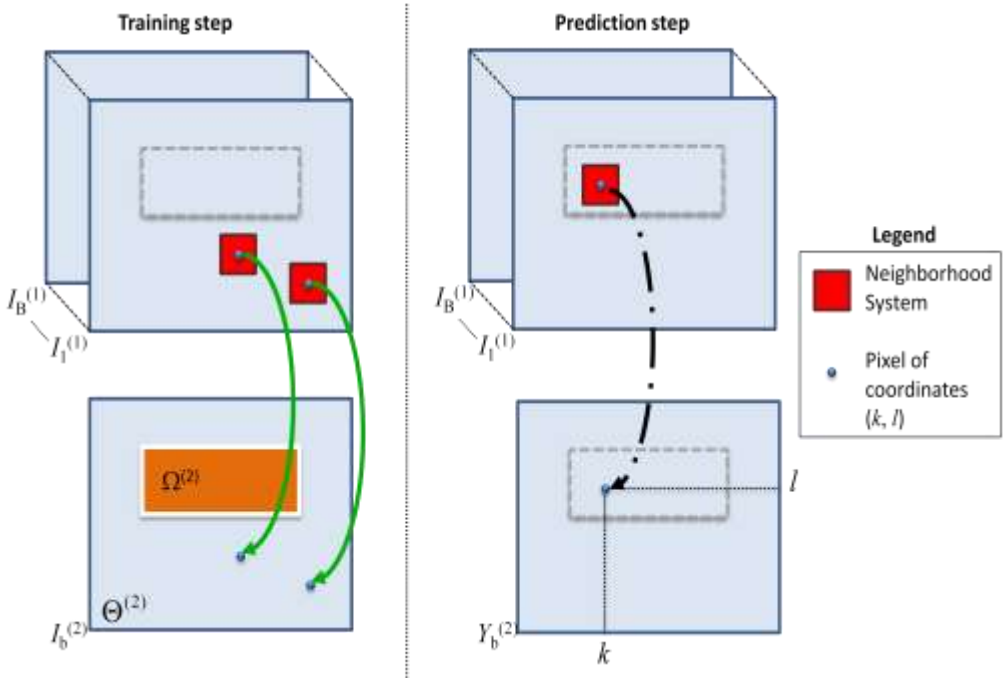


Figure 4.1. Example of the reconstruction process: training and prediction steps.

4.3. Proposed Solution

4.3.1. ε-insensitive support vector regression

In any regression problem, one seeks a function that best links the input to the output spaces. ε-SVR is a prediction approach, which takes origin from the statistical learning theory ([13], [14]) and has proved to be efficient in different contexts including the remote sensing field [12]. It aims at getting a function $f(\boldsymbol{x})$ that has at most ε deviation from the desired target y and at the same time is as smooth as possible. This can be obtained by performing a non-linear projection, mapping the data from the original d-dimensional domain to a higher dimensional feature space, i.e., $\Phi(\boldsymbol{x}) \in \mathbb{R}^{d'}$ $(d' > d)$, in order to increase the flatness of the function, and accordingly to approximate it in a linear way as:

$$f(\boldsymbol{x}) = \boldsymbol{w}^* \cdot \Phi(\boldsymbol{x}) + b^* . \tag{4.2}$$

The optimal hyperplane defined by the weight vector $\boldsymbol{w}^* \in \mathbb{R}^{d'}$ and the bias $b^* \in \mathbb{R}$ is the one that minimizes the cost function:

$$\Psi(\boldsymbol{w}, \xi) = \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*). \tag{4.3}$$

and is subject to the following constraints:

$$\begin{cases} y_i - (\mathbf{w} \cdot \Phi(\boldsymbol{x}_i) + b) \leq \varepsilon + \xi_i \\ (\mathbf{w} \cdot \Phi(\boldsymbol{x}_i) + b) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}, \tag{4.4}$$

where ξ_i and ξ_i^* are the slack variables introduced for the samples which are not in the ε-tube, depending on whether they lie above or below the tube, respectively. The constant C represents a regularization parameter with which it is possible to obtain a compromise between the flatness (model complexity) of the function $f(\boldsymbol{x})$ and the accuracy of the regression (training error). Note that the sum in (4.3) takes into account all the available N training samples [15].

There exists an optimization solution which can reformulate and solve the previous problem by using the Lagrange multipliers and a quadratic programming solver, leading to the following final prediction model:

$$f(\boldsymbol{x}) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) \cdot K(\boldsymbol{x}_i, \boldsymbol{x}) + b^* , \tag{4.5}$$

where $K(\cdot,\cdot)$ is a kernel function that characterizes the dot product behavior in a feature space defined by $\Phi(\cdot)$ and only samples with nonzero coefficients α_i and α_i^* are the SVs which can lie on the ε-tube and contribute to the prediction. The function kernel $K(\cdot,\cdot)$ must satisfy the Mercer's theorem [12]. In the following subsection, we will briefly describe the most common kernel functions [16].

4.3.2. Common kernel functions

A simple kernel function is the linear (LIN) kernel, which is defined as:

$$K_{LIN}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \boldsymbol{x_1}^T \boldsymbol{x_2} + c, \tag{4.6}$$

where c is a constant.

Another kernel is the non-stationary polynomial (POL) kernel:

$$K_{POL}(\boldsymbol{x_1}, \boldsymbol{x_2}) = (\beta \boldsymbol{x_1}^T \boldsymbol{x_2} + c)^d, \tag{4.7}$$

where β is a parameter for the slope, c is a constant and d the degree of the polynomial.

Another common and most versatile kernel is the Gaussian function, also known as radial basis function (RBF) kernel:

$$K_{RFB}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \exp(-\gamma \|\boldsymbol{x_1} - \boldsymbol{x_2}\|^2), \tag{4.8}$$

where $\gamma = 1/2\sigma^2 > 0$ is a width parameter.

Generally, all kernel parameters play an important role in the performance of the kernel and thus must be carefully tuned jointly with the regularization parameter C and precision parameter ε (e.g., by adopting a grid search via a n-fold cross-validation procedure, CV).

4.3.3. Kernel function

In our reconstruction problem, since we deal with heterogeneous features (i.e., position and radiometric information), the proposed idea is to design a fusion kernel derived from a combination of single kernels, each for every feature typology. From [17], if a set of kernels forms a convex cone, closed and under pointwise convergence, or in other words, if K_1 and K_2 are valid kernels, and $\beta_1, \beta_2 \geq 0$, then also the following expression is a valid kernel:

$$\beta_1 K_1 + \beta_2 K_2 \, . \tag{4.9}$$

It is possible to rewrite this expression, using only one weight constant to balance the sum:

$$\mu K_1 + (1 - \mu) K_2 \, . \tag{4.10}$$

This kernel combination opens the way to design a large number of new kernels, by linearly combining the most common kernels (e.g., LIN-LIN, LIN-POL, POL-LIN, LIN-RBF, etc.). Because of space limitations, we will just consider the combination based on two RBF kernels, found empirically to be the best one:

$$K_{RBF-RBF} = \mu K_{RBF1} + (1 - \mu) K_{RBF2} \, , \tag{4.11}$$

where $0 \leq \mu \leq 1$ and each kernel has its proper kernel width γ value, which is evaluated with a grid search CV. Note that, we also try to add all the spatial information in a dedicate kernel; in another simulation we add a third kernel K_3 to have a specific kernel for each kind of information (spatial coordinates, radiometric and neighborhood). In both cases, we do not achieve interesting changes in the results.

In the following, we will refer to the method which applies this kind of kernel, as "Gaussian kernel combination regression" (GKCR).

4.3.4. Feature Vector

The contextual nonlinear prediction (CNP) was originally defined to adopt a SVR for the reconstruction of cloudy regions in multitemporal images, by exploiting a "local neighborhood" [7]. In our implementation, we propose to add position information (i.e., pixel coordinates (k, l) in Figure 4.1) in a specific kernel K_1 (see Figure 4.2). The new addition of this information provides to

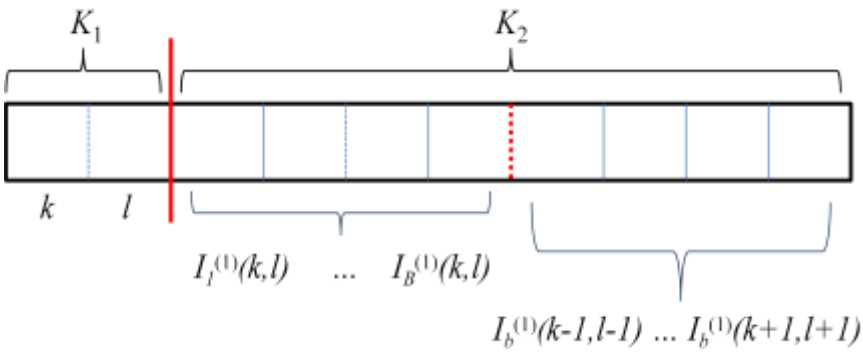


Figure 4.2. Structure of the proposed feature vector: the first 2 elements correspond to the coordinates of the pixel, the next B refer to the corresponding pixel in all the spectral bands of $I^{(1)}$ and the last four to the neighborhood of the pixel of interest in $I_b^{(1)}$ where b represents the spectral band that is targeted.

restrict locally the sample similarity. It induces a trend on a regular sampling in the selection of the SVs. The second kernel K_2 considers radiometric information, composing the second part of the feature vector. Actually, this part can be subdivided into two subparts. The first B values come from the spectral information extracted from image $I^{(1)}$. The last four values come from the spatial information conveyed in a spatial neighborhood system of first order and centered on the pixel of interest in the image $I_b^{(1)}$where b represents the band that is targeted. Note that the pixel of interest comes from $\Theta^{(1)}$ or $\Omega^{(1)}$ depending if training or test step is considered, respectively.

4.4. Experimental Results

4.4.1. Dataset description

To cope with this missing data problem, at least two images are necessary. In our simulations, we will assume that just one of them can contain cloudy regions. The test images we used come from the optical high resolution FORMOSAT-2 satellite [18]. The images represent part of the *Arcachon* basin in the south region of Aquitaine, in France. Images convey 4 spectral bands (blue, green, red and near-infrared) and are characterized by a spatial resolution of 2 meter. They were acquired in two different days. Image $I^{(1)}$ was taken on the 24th of June 2009, while image $I^{(2)}$ was taken three weeks later, on the 16th of July. The two images do not contain significant differences in the spatial structures. For our experiments, the two images have been cut out into 400×400 pixel-size images containing different land cover typologies, principally grass and urban areas (see the region of interest ROI A and B in Figure 4.4 and Figure 4.6, respectively). We artificially obscured part of $I^{(2)}$ image, as shown in Figure 4.4(b) and Figure 4.6(b). The resulting mask makes it possible to quantify the reconstruction accuracy by comparing the reconstruction result with the true pixel values. As an accuracy measure, we adopt the well-known peak signal-to-noise ratio (PSNR) measure [19] evaluated on the images, for the training and test steps.

4.4.2. Experiments

Before the parameter estimation phase, we decide to fix the number of training samples (#TR) to adopt in the regression, in correlation with the number of missing values, i.e., the test samples (#TS). In particular, we took three different ratios #TR/#TS, which are: $1/5$, $1/3$, and $3/4$, where the number of test samples $\#\mathrm{TS} = 2060$. Note that the training samples are collected adopting 10 times shifted sampling grids (see an example of sampling in Figure 4.4(a) and in Figure 4.6(a)).

Before starting with the experiments, all different kinds of data represented in the feature vector have to be normalized in the interval [0,1]. Afterwards, one of the first evaluations is to find which combination of RBF kernels returns the better result, namely which value of μ to adopt in Eq. (4.11). The answer in this case comes from empirical estimations. We apply for each value of μ a complete grid search, finding the best parameters (with $C \in \{10^{-3}, 2\cdot10^{-3}, 5\cdot10^{-3}, \cdots, 10^{2}\}$, $\gamma_i \in \{10^{-4}, 2\cdot10^{-4}, 5\cdot10^{-4}, \cdots, 1\}$ and $\varepsilon \in \{0.01, 0.02, \cdots, 0.1\}$), considering the first of the three different samples ratios (i.e., $\#\mathrm{TR}/\#\mathrm{TS} = 1/5$). From an analogous problem, where an RBF kernel was adopted (see the CNP in [7], similar to our kernel with $\mu = 0$ in Eq. (4.11)), we start to weight more K_1, increasing μ value. In other words, we initiate favoring radiometric information over spatial coordinates, until we reach $\mu = 1$, where the radiometric information disappears (see this evaluation in Table 4.I and in Figure 4.3). Note that if we take into account only one of the two parts of the RBF kernels we obtain worst reconstructions. For example if $\mu = 0$, we have a lack of information, and when $\mu = 1$ the estimation does not take into account the radiometric information; it considers only the localization of the pixel. It is worth noting the result seems to be more stable between $\mu = 0.1$ to $\mu = 0.9$. Similar results are obtained for ROI B and for the three different ratios, where the best result still arises with $\mu = 0.1$.

TABLE 4.I
QUANTITATIVE RESULTS OBTAINED BY THE GKCR IN THE TWO ROI, INCREASING THE VALUE OF μ AND ADOPTING #TR/#TS=1/5 RATIO.

	ROI A					**ROI B**				
μ	ε	C	γ_1	γ_2	**PSNR**	ε	C	γ_1	γ_2	**PSNR**
0	0.05	100	-	0.01	29.88	0.05	500	-	0.05	25.13
0.1	0.05	10	0.001	0.1	**30.73**	0.05	500	0.001	0.02	**26.09**
0.2	0.05	10	0.001	0.1	30.68	0.05	500	0.001	0.02	26.01
0.3	0.05	5	0.001	0.2	30.65	0.05	500	0.001	0.02	25.98
0.4	0.05	2	0.001	0.5	30.58	0.05	500	0.001	0.02	25.85
0.5	0.05	2	0.001	0.5	30.55	0.05	200	0.001	0.05	25.60
0.6	0.05	2	0.001	1	30.51	0.05	200	0.001	0.05	25.50
0.7	0.05	2	0.001	1	30.46	0.05	500	0.001	0.05	25.37
0.8	0.05	5	0.001	1	30.41	0.05	500	0.001	0.05	25.31
0.9	0.05	5	0.001	1	30.28	0.05	200	0.001	0.1	25.21
1	0.05	500	0.2	-	22.29	0.05	500	0.1	-	17.52

Once that μ value is found, the corresponding SVR parameters for both kernels, K_1 and K_2 are known: the regularization parameter *C*, the two width parameters γ_i, and the precision parameter ε. A first important result which comes out in all the experimentations is that ε does not change

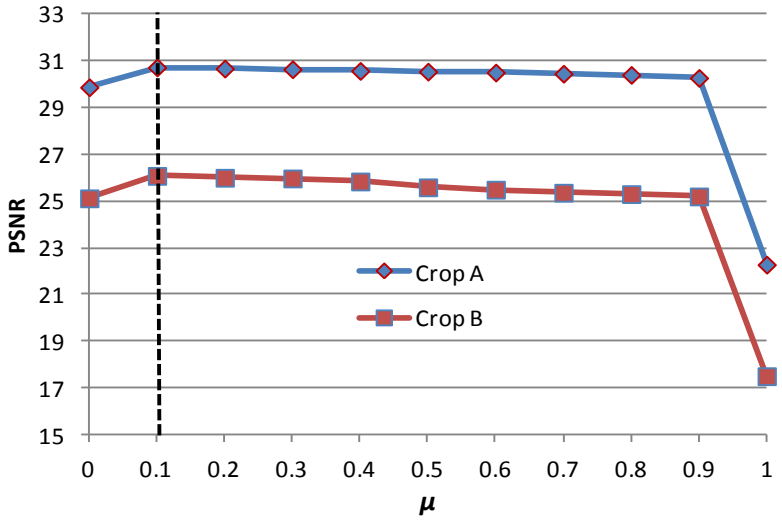


Figure 4.3. Evolution of training and test PSNR values varying μ parameter, for the two crop regions, adopting #TR/#TS=1/5 ratio.

significantly the CV results; for this reason we decide to fix its value for the following simulations at $\varepsilon = 0.05$, in order to make the grid search faster. Having fixed μ and ε values, it comes out that generally $\gamma_1 \ll \gamma_2$. This last means that the location kernel K_1 is less sensitive than the radiometric kernel K_2, but both are still important for achieving a good reconstruction.

In the following we present qualitatively results obtained applying $\#TR/\#TS = 1/5$ ratio, which means to adopt 395 training samples to reconstruct 2060 test values. Figure 4.4(a) shows ROI A of image $I^{(1)}$ with the TR samples marked with yellow points, while Figure 4.4(c) gives the corresponding reconstructed image, obtained with only 19 SVs (highlighted in yellow) and a high accuracy value (PSNR=30.72 dB in average). Given the fact that the missing area represents an urban region, we note that almost all the SVs come from similar regions. Similarly, Figure 4.6(a) shows ROI B of image $I^{(1)}$ with the 395 TR samples marked with yellow points, and Figure 4.6(c) the reconstruction of the missing area with the position of the selected SVs. In this case, only half of the SVs belong to the urban region. This result is probably due to the fact that the dominating class area in ROI B is the vegetation. It may explain the lower value of the PSNR (PSNR=26.91 dB in average) and the need of more SVs (#SV=34). For more visual details about the reconstructions of the two ROI images, see the zoom images in Figures 4.5 and 4.7.

(a) (b) (c)

Figure 4.4. ROI A: (a) $I^{(1)}$ image with highlighted TR samples (#TR = 395 samples), (b) $I^{(2)}$ image with the addition of the mask, and (c) reconstructed $I^{(2)}$ image with highlighted SVs.

(a) (b) (c)

Figure 4.5. Zoom of area of interest in Figure 4.4: (a) $I^{(1)}$ image, (b) $I^{(2)}$ original image, and (c) $I^{(2)}$ reconstructed image.

(a) (b) (c)

Figure 4.6. ROI B: (a) $I^{(1)}$ image with highlighted TR samples (#TR = 395 samples), (b) $I^{(2)}$ image with the addition of the mask, and (c) reconstructed $I^{(2)}$ image with highlighted SVs.

(a) (b) (c)

Figure 4.7. Zoom of area of interest in Figure 4.6: (a) $I^{(1)}$ image, (b) $I^{(2)}$ original image, and (c) $I^{(2)}$ reconstructed image.

4.4.3. Comparative analysis

In this subsection, we report a comparison with state-of-the-art reconstruction techniques, where the experimental settings are in general similar as for the GKCR method. The first one is a multiresolution inpainting (MRI) technique which does not need temporal information to reconstruct the missing region [11]. It adopts 3×3, 6×6 and 9×9 square patches. In similar way

as GKCR, other techniques perform regression within a multitemporal reconstruction context. The first one is the contextual multiple linear prediction (CMLP) which exploits only temporal information [7]. A second one is the contextual nonlinear prediction (CNP) which exploits local-spectral information adopting in a local neighborhood of the missing area [7]. A third one is the SM1 contextual spatiospectral postreconstruction (CSSPR), which starting from CMLP, takes advantage of the local properties in a predefined neighborhood system (i.e., 3×3) [8]. CMLP adopts a simple linear least squares regression, whereas the last two techniques exploit RBF kernels in a SV regression, where similar CV step is adopted. Table 4.II lists the results of the investigated reconstruction techniques carried on ROI A (mean and standard deviation over 10 runs). The poorest results are yielded by the MRI and the CMLP methods. They return unsatisfactory results; MRI due to its unsupervised nature and to the fact it does not make use of the temporal dimension of the image sequence. Even worse is the second, the CMLP, mainly because it adopts a simple linear regression. The best method is the GKCR, which exhibits the higher PSNR values. We can perceive that the accuracy results are stable in the three ratios, adopting different sizes of training sets. Furthermore we observe that from the three ratio cases, it is still possible to reach high accuracy also with a reduced complexity, namely a smaller number of SVs, especially if compared with CNP method. Its underlying idea to exploit all available information (temporal, spectral and spatial information) within a suitable kernel fusion framework has provided it with a superior capability to handle the reconstruction issue.

TABLE 4.II
COMPARISON BETWEEN THE 5 INVESTIGATED RECONSTRUCTION TECHNIQUES FOR ROI A.

	#TR	$\frac{\#TR}{\#TS}$	**Time [s]**	**Average PSNR**	**Stand. dev. PSNR**	**#SV**
MRI	--	--	67	22.48	--	--
CMLP	395	1/5	1	20.94	0.70	--
	723	1/3	1	20.93	0.51	--
	1580	3/4	1	20.98	0.15	--
CNP	395	1/5	17	29.35	0.32	29
	723	1/3	51	29.35	0.48	56
	1580	3/4	189	29.52	0.17	122
CSSPR	395	1/5	17	29.23	0.32	**19**
	723	1/3	52	29.33	0.34	33
	1580	3/4	197	29.54	0.15	63
GKCR	395	1/5	20	**30.72**	0.54	**19**
	723	1/3	62	**30.96**	0.42	**32**
	1580	3/4	244	**31.22**	0.27	**61**

4.5. Conclusion

In this chapter, we have proposed a new reconstruction method for missing data due to cloud covers. It integrates in the reconstruction process two types of information: 1) the position of the missing value and 2) the radiometric information. Their fusion performed by means of a kernel combination together with the power of the support vector regression have made it particularly promising as suggested by the experiments. The price of its superior effectiveness is however a

higher (but still contained) computation time because of the larger number of free parameters to estimate compared to reference reconstruction methods.

4.6. Acknowledgment

The authors would like to thank C. C. Chang and C. J. Lin for supplying the software library LIBSVM used in this research. The authors are grateful to CNES for making available online the FORMOSAT-2 images.

4.7. References

[1] X. Zhang, F. Qin, and Y. Qin, "Study on the thick cloud removal method based on multi-temporal remote sensing images," in *Proc. ICMT*, Malaga, Spain, Jul. 2010, pp. 1–3.

[2] B. A. Latif, R. Lecerf, and G. Mercier, "Preprocessing of low-resolution time series contaminated by clouds and shadow," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2083–2096, Jul. 2008.

[3] D-C. Tseng, H-T. Tseng and C-L. Chien, "Automatic cloud removal from multi-temporal SPOT images," *J. of Applied Mathematics and Computation*, vol. 205, pp. 584–600, Nov. 2008.

[4] S. C. Liew, M. Li, and L. K. Kwoh, "Automated production of cloud-free and cloud-shadow image mosaics from cloudy satellite imagery," in *Proc. XXth ISPRS*, Istanbul, Turkey, Jul. 2004, pp. 523–530.

[5] E. H. Helmer and B. Ruefenacht, "Cloud-free satellite image mosaics with regression trees and histogram matching," *J. of the American Society for Photogrammetry and Remote Sensing*, vol. 71, pp. 1079–1089, Sep. 2005.

[6] S. Lee and M.M. Crawford, "An adaptive reconstruction system for spatially correlated multispectral multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 4, pp. 494–508, Jul. 1991.

[7] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 442–455, Feb. 2006.

[8] S. Benabdelkader and F. Melgani, "Contextual spatiospectral postreconstruction of cloud-contaminated images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 204–208, Apr. 2008.

[9] N. T. Hoan and R. Tateishi, "Cloud removal of optical image using SAR Data for ALOS applications. Experimenting on simulated ALOS data," in *Proc. ISPRS*, Beijing, China, Jul. 2008, vol. 37, pp. 379–384.

[10] A. Maalouf, P. Carré, B. Augereau and C. F. Maloigne, "A bandelet-based inpainting technique for clouds removal from remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2363–2371, Jul. 2009.

[11] L. Lorenzi, F. Melgani and G. Mercier, "Inpainting strategies for reconstruction of missing data in VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 914–918, Sep. 2011.

[12] G. Camp-Valls and L. Bruzzone, *Kernel methods for remote sensing data analysis*, John Wiley and Sons, 2009.

[13] V. Vapnik, *The nature of statistical learning theory*. New York: Springer-Verlag, 1995.

[14] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[15] A. Smola and B. Schölkopf, *A tutorial on support vector regression*. London, U.K.: Roy. Holloway College, Univ. London, 1998.

[16] M. G. Genton, "Classes of kernels for machine learning: A statistics perspective," *J. of Machine Learning Research*, vol. 2, pp. 299–312, Dec. 2001.

[17] B. Schölkopf and A. J. Smola, *Learning with kernels*. Cambridge, U.K.: The MIT Press, 2002.

[18] C-C. Liu, "Processing of FORMOSAT-2 daily revisit imagery for site surveillance," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3206–3215, Nov. 2006.

[19] A.K. Jain, *Fundamentals of digital image processing*. New York: Prentice Hall, 1988.

5. Missing Area Reconstruction in Multispectral Images Under a Compressive Sensing Perspective

Abstract *– The intent of this chapter is to propose new methods for the reconstruction of areas obscured by clouds. They are based on compressive sensing (CS) theory, which allows to find sparse signal representations in underdetermined linear equation systems. In particular, two common CS solutions are adopted for our reconstruction problem: the basis pursuit (BP) and the orthogonal matching pursuit (OMP) methods. A novel alternative CS solution is also proposed through a formulation within a multiobjective genetic optimization scheme. To illustrate the performances of the proposed methods, a through experimental analysis on FORMOSAT-2 and SPOT-5 multispectral images is reported and discussed. It includes a detailed simulation study that aims at assessing the accuracy of the methods in different qualitative and quantitative cloud-contamination conditions. Compared with state-of-the-art techniques for cloud removal, the proposed methods show a clear superiority, which makes them a promising tool in cleaning images in the presence of clouds.*

The work presented in this chapter has been submitted to *IEEE Trans. Geosci. Remote Sens.*; Co-authors: F. Melgani, G. Mercier.

5.1. Introduction

Depending on the end-user requirements, clouds in remotely sensed imagery may or not represent an unwanted source of noise. In case they are viewed as a noise source, several methodologies have been developed in the past in order to cope with this problem. Generally, the common approach first detects the contaminated regions, and in a second instance, attempts to remove the clouds by substituting them with cloud-free estimations. In this paper, we will focus on the second part of the approach. In the related literature, one of the first proposed techniques produces a cloud-free image mosaics by compositing several cloudy optical satellite images acquired from the same area [1]. Other techniques reconstruct missing areas by exploiting the inpainting approach. For instance, a recent work suggests to enrich the region (patch) search process by including local image properties or by isometric transformations, or to reformulate the problem under a multiresolution processing scheme [2]. Another synthesis technique is proposed by Maalouf *et al.* [3] where information is propagated by a technique based on the bandelet transform, a special case of the wavelet transform. Accurate reconstructions may be obtained through prediction techniques, if temporal information is available. A first work exploiting the temporal prediction principle was presented in [4], where a least-squares linear prediction with escalator structure is implemented. Unsupervised contextual prediction was also adopted in [5], where local spectro-temporal relationships are used to predict missing data through linear or nonlinear regression. Tseng *et al.* [6] proposed a method to correct radiometric inconsistencies of cloud-contaminated images and their corresponding temporal images by generating a cloud-free mosaic image for a multitemporal SPOT dataset. In order to ameliorate the transition between two mosaic parts, a wavelet-based fusion is adopted. More recently, a cloud removal method, based on information cloning was developed [7]. The authors propose to clone cloud-free information from a set of multitemporal images, adopting a patch-based reconstruction method formulated as a Poisson equation and solved using a global optimization process.

Recently, compressive sensing (CS), also known as compressive sampling or sparse representation, was introduced by Donoho [8] and Candès [9]. CS theory aims at recovering an unknown sparse signal from a small set of linear projections. By exploiting this new and important result, it is possible to obtain equivalent or better representations by using less information compared with traditional methods (i.e., lower sampling rate or smaller data size) [10]. CS has proved to be a powerful tool for several applications, such as: acquisition, representation, regularization in inverse problem, feature extraction and compression of high-dimensional signals, and applied in different research fields: signal processing, object recognition, data mining, and bioinformatics [11]. In these fields, CS has been adopted to cope with several tasks: face recognition [12], image super-resolution [13], segmentation [14], denoising [15], inpainting [16] and classification [17]. Note that images are a special case of signals which hold a natural sparse representation, with respect to fixed bases, also called dictionary (i.e.: Fourier, wavelet) [18]. In the literature, a common dictionary choice is the one proposed in [11], the K-SVD (K-means Singular Value Decomposition) may be a solution. It adapts dictionaries by iteratively alternating between sparse coding of exemplar-based on the current dictionary and a process of updating the dictionary atoms to better fit the data. But to simplify the selection of a proper dictionary, in some cases the training pixels or patches (which are supposed to be known) are collected directly from the image of

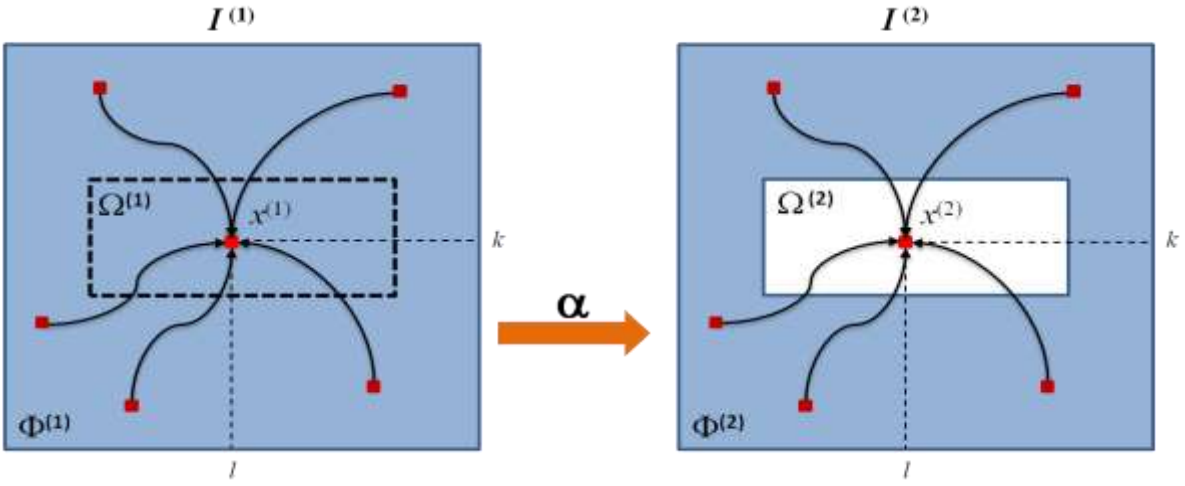
interest ([12], [16]), or by random sampling the training area [19]. More recently, CS theory has been applied also in the remote sensing (RS) field. For example, there exist specific applications such as: 1) for image pan-sharpening as in [19], where CS theory ensures that a compressible signal can be correctly recovered from global linear sampled data; 2) for hyperspectral image classification [20], where two approaches have been proposed: i) in the first one, an explicit smoothing constraint is imposed by forcing the Laplacian vector to zero; and ii) the second approach relies on a joint sparsity model where hyperspectral pixels are simultaneously represented by linear combinations; 3) for super-resolution restoration [21], where high resolution image patches are recovered from the down-sampled low-resolution ones; and 4) for saving camera size, power consumption and storage burden during the remote sensing image acquisition process [22].

In this paper, we propose three novel methods to solve the problem of the reconstruction of missing data due to the presence of clouds. Given a cloud-free and a cloud-contaminated image, each of the missing measurements is recovered applying the CS theory in which cloud-free pixels are exploited. More in details, we first apply two of the most common CS methods which approximate the CS solution, namely the basis pursuit (BP) and the orthogonal matching pursuit (OMP). Furthermore, we propose an alternative CS solution, which exploits the search capabilities of genetic algorithms (GAs). Experimental results on two temporal multispectral images acquired by FORMOSAT-2 and SPOT-5 optical satellites are reported and discussed. They include a thorough simulation to assess accurately the performances of the three proposed methods and other two cloud removal techniques in terms of reconstruction quality.

This Chapter is organized as follows. In next section, we formulate the cloud-contaminated image reconstruction problem. In Section 5.3, we introduce the CS theory and describe two common CS methods, whereas the GA-based method is described in Section 5.4. Experimental results are illustrated in Section 5.5. Section 5.6 draws the conclusions.

5.2. Problem Formulation

Let us consider two multispectral (*B*-channels) images $I^{(1)}$ and $I^{(2)}$ acquired by an optical sensor at two different dates and registered over the same geographical area. Let us suppose that the two acquisitions are temporally close to each other and thus the images are characterized by a similar spatial structure. The objective of the proposed methods is to reconstruct any area contaminated by clouds. We remark that the cloud detection (i.e., finding the position of clouds in an image) is out of the scope of this paper. At this level, we make the hypothesis that image $I^{(2)}$ is obscured by the presence of clouds and a cloud/non-cloud binary classifier has been adopted to discriminate between these two classes. We will call this cloudy area in the image $I^{(2)}$ as target region $\boldsymbol{\Omega}^{(2)}$ and the remaining part as source region $\boldsymbol{\Phi}^{(2)}$, subject to $I^{(2)} = \boldsymbol{\Omega}^{(2)} \cup \boldsymbol{\Phi}^{(2)}$ (following classical notation in the inpainting literature). Image $I^{(1)} = \boldsymbol{\Omega}^{(1)} \cup \boldsymbol{\Phi}^{(1)}$ does not contain clouds since it is supposed cloud-free. Our aim is to generate a new image $I^{(2)}$ without clouds.



(a) (b)

Figure 5.1. Illustration of the reconstruction principle under a CS perspective.

We will assume that any pixel $x^{(1)} \in \boldsymbol{\Omega}^{(1)}$ can be expressed as a linear combination of pixels in region $\boldsymbol{\Phi}^{(1)}$ of $I^{(1)}$ (see Figure 1(a)). In other words, in $I^{(1)}$, we have:

$$x^{(1)} = \boldsymbol{\Phi}^{(1)} \cdot \boldsymbol{\alpha}, \quad \forall x^{(1)} \in \boldsymbol{\Omega}^{(1)}, \tag{5.1}$$

where $\boldsymbol{\alpha}$ is an unknown weight vector associated with the considered pixel $x^{(1)}$ and having the same dimension as the number of pixels belonging to $\boldsymbol{\Phi}^{(1)}$. The problem at this point is to infer $\boldsymbol{\alpha} = f\left(\boldsymbol{\Phi}^{(1)}, x^{(1)}\right)$. Once $\boldsymbol{\alpha}$ is computed, if we assume that $I^{(1)}$ and $I^{(2)}$ are temporally close, so that the scene did not change in-between the two observations, it will be possible to reuse the $\boldsymbol{\alpha}$ coefficients to reconstruct the spatially corresponding pixel in the missing area $\boldsymbol{\Omega}^{(2)}$, adopting the previous formulation for $I^{(2)}$, i.e., $\hat{x}^{(2)} = \boldsymbol{\Phi}^{(2)} \cdot \boldsymbol{\alpha}$ (see Figure 5.1(b)). In other words, for each pixel $x^{(1)} \in \boldsymbol{\Omega}^{(1)}$, we evaluate $\boldsymbol{\alpha}$, and in a second moment we reuse this weight vector to return an estimation of $x^{(2)} \in \boldsymbol{\Omega}^{(2)}$:

$$\begin{array}{l} \text{From } I^{(1)}: \quad \boldsymbol{\alpha} = f\left(\boldsymbol{\Phi}^{(1)}, x^{(1)}\right) \\ \text{To } I^{(2)}: \quad \hat{x}^{(2)} = \boldsymbol{\Phi}^{(2)} \cdot \boldsymbol{\alpha} \end{array}, \tag{5.2}$$

where $f(\cdot)$ represents an estimation function. Finally, we note that, differently from reference techniques as the one presented in [5], all image channels are processed simultaneously. In the next two sections, different methods are proposed to solve the issue of the estimation of the weight vectors.

5.3. Reconstruction via Compressive Sensing

5.3.1. Generalities on Compressive Sensing

Compressive sensing (CS) is a useful way to obtain a sparse representation of a signal. It relies on the idea to exploit redundancy (if any) in the signals [8]-[9]. Usually signals like images are sparse, as they contain, in some representation domain, many coefficients close to or equal to zero. CS starts taking a weighted linear combination of pixels in a basis in which the signal is

assumed to be sparse. The fundamental of the CS theory is the ability to recover with relatively few measurements $x = D \cdot \boldsymbol{\alpha}$ by solving the following L_0-minimization problem:

$$min\|\boldsymbol{\alpha}\|_0 \text{ subject to } x = D \cdot \boldsymbol{\alpha}, \tag{5.3}$$

where D is a dictionary with a certain number of atoms, x is the original signal which can be represented as a sparse linear combination of these atoms, and the minimization of $\|\cdot\|_0$, the L_0-norm, corresponds to the maximization of the number zeros in $\boldsymbol{\alpha}$, following this formulation: $\|\boldsymbol{\alpha}\|_0 = \#\{i: \alpha_i \neq 0\}$. Equation (5.3) represents a NP-hard problem, which means that it is computationally infeasible to solve. Following the discussion of Candès and Tao [23], it is possible to simplify the evaluation of (5.3) in a relatively easily linear programming solution. They demonstrate that, under some reasonable assumptions, minimizing L_1-norm is equivalent to minimizing L_0-norm, which is defined as $\|\boldsymbol{\alpha}\|_1 = \sum_i |\alpha_i|$. Accordingly, it is possible to rewrite Eq. (5.3) as:

$$\min\|\boldsymbol{\alpha}\|_1 \quad \text{subject to} \quad x = D \cdot \boldsymbol{\alpha}. \tag{5.4}$$

In the literature there exist several algorithms for solving optimization problems similar to the one expressed in Eq. (5.4). In the next subsections, we briefly introduce two of them, which represent the most common solutions from the literature.

5.3.1.1. CS Solutions

1) Basis Pursuit (BP)

A well-known solution for problem Eq. (5.4) is the basis pursuit (BP) principle [23]-[24]. It suggests a convexification of the problem by using the L_1-norm instead of L_0. This means that the best approximation of the problem becomes equal to a support minimization problem. BP finds signal representations in overcomplete dictionaries by convex, nonquadratic optimization technique, solving problem Eq. (5.4). It can be reformulated as a linear programming (LP) problem, and solved using modern interior-point methods, simplex methods, or other techniques, such as homotopy techniques [25]. Given that, it is possible to rewrite the L_1-norm in Eq. (5.4) as:

$$\|\boldsymbol{\alpha}\|_1 = \sum_i |\alpha_i| = \sum_i u_i + v_i \,, \tag{5.5}$$

$$\text{where } \begin{cases} \alpha_i = u_i,\ v_i = 0 \quad if\ \alpha_i \geq 0 \\ \alpha_i = -v_i,\ u_i = 0\ if\ \alpha_i \leq 0 \end{cases}.$$

Substituting it in Eq. (5.4), it allows to perform a linear minimization problem. Note that, if the original signal x is sufficiently sparse, the recovery via BP is provably exact.

2) Orthogonal Matching Pursuit (OMP)

One of the easiest and fastest alternative technique is the orthogonal matching pursuit (OMP), an improved version of the matching pursuit (MP) method. MP finds the atom that has the highest correlation with the signal. It subtracts off the correlated part from the signal and then iterates the procedure on the resulting residual signal [26]-[27]. The algorithm approximates the signal *x*, considering these two decompositions [24]:

$$x = \sum_{d \in D} \alpha_d \varphi_d = \sum_{i=1}^{m} \alpha_{d_i} \varphi_{d_i} + R^{(m)} \,, \tag{5.6}$$

where dictionary *D* is a collection of atom vectors $\{\varphi_d\}_{d\in D}$ and $R^{(m)}$ is a residual. Starting from an initial approximation $x^{(0)} = 0$ and residual $R^{(0)} = x$, it builds up a sequence of sparse approximations stepwise. At stage *k*, it identifies the dictionary atom that best correlates with the residual and then adds to the current approximation a scalar multiple of that atom, so that $x^{(k)} = x^{(k-1)} + \alpha_k \varphi_{d_k}$, where $\alpha_k = \langle R^{(k-1)}, \varphi_{d_k} \rangle$ and $R^{(k)} = x - x^{(k)}$. After m steps, one has a representation of the form Eq. (5.6), with residual $R = R^{(m)}$, where the original signal x is decomposed into a sum of dictionary elements, that are chosen to best match its residues.

Unfortunately the convergence speed of this algorithm is not fast. To overcome this drawback, an improved solution called orthogonal MP (OMP) was developed. Differently from MP, OMP updates the coefficients of the selected atoms at each iteration so that the resulting residual vectors are orthogonal to the subspace spanned by the selected atoms. When stopped after only few iterations, it generally yields a satisfactory approximation, using only few atoms [26]-[27].

3) BP vs. OMP

From the literature [28]-[29], it comes out that BP and OMP algorithms provide in general good performances in reconstruction problems. Nonetheless, BP is considered more powerful than OMP, since it can recover with high probability all sparse signals and is more stable. On the contrary, OMP results attractive for its fast convergence and in its ease of implementation.

5.4. Genetic Algorithm

5.4.1. General Concepts on GA

Genetic algorithms (GA) are a part of evolutionary computation which solves optimization problem by mimicking the principles of biological evolution [30]-[31]. A genetic optimization algorithm performs a search by regenerating a population of candidate solutions (or individuals) represented by chromosomes. From one generation to the next, the population is improved following biological rules, adopting deterministic and nondeterministic genetic operators. In general, a common GA involves the following steps. First, an initial population of chromosomes is randomly generated. Then, the goodness of each chromosome is evaluated according to a predefined fitness function representing the aim of the optimization. Evaluating the fitness function allows to keep or discard chromosomes, by using a proper rule based on the principle that the better the fitness, the higher the chance of being selected. Once the selection of the best chromosomes is done, the next step is devoted to the reproduction of a new population. This is done by genetic operators such as crossover and mutation operators. All these steps are iterated until some predefined condition is satisfied (e.g., maximum number of generations, or fitness value limit).

Several multiobjective GA-based approaches have been proposed in the literature [32], such as SPEA-II [33], PAES [34] and NSGA-II [35]. In this paper, we will adopt the nondominated sorting solution (NSGA-II) for its low computational requirements, its aptitude to distribute uniformly the optimal solutions along the Pareto front [35] and its successful application to different remote sensing problems [36]-[38]. NSGA-II is based on the concept of nondominance. A solution

s_1 is said to dominate another solution s_2, if s_1 is not worse than s_2 in all objectives and better than s_2 in at least one objective. A solution is said to be nondominated if it is not dominated by any other solution. Figure 5.2 illustrates a multiobjective GA chromosome selection problem with two fitness functions f_1 and f_2. As GA does, NSGA-II starts by generating a random parent population. Individuals (chromosomes) selected through a crowded tournament selection undergo crossover and mutation operations to form an offspring population. Both offspring and parent population are then combined and sorted into fronts of decreasing dominance (rank). After the sorting process, the new population is filled with solutions of different fronts starting from the best one. If a front can only partially fill the next generation, crowded tournament selection is applied again to ensure diversity. Once the next generation population has been completely filled, the algorithm loops back to create a new offspring population and the process continues up to convergence.

5.4.2. GA setup

The design of a multiobjective GA optimization relies upon two components, the chromosome structure and the fitness functions, which encode the considered optimization problem and show the direction to obtain the best solution, respectively.

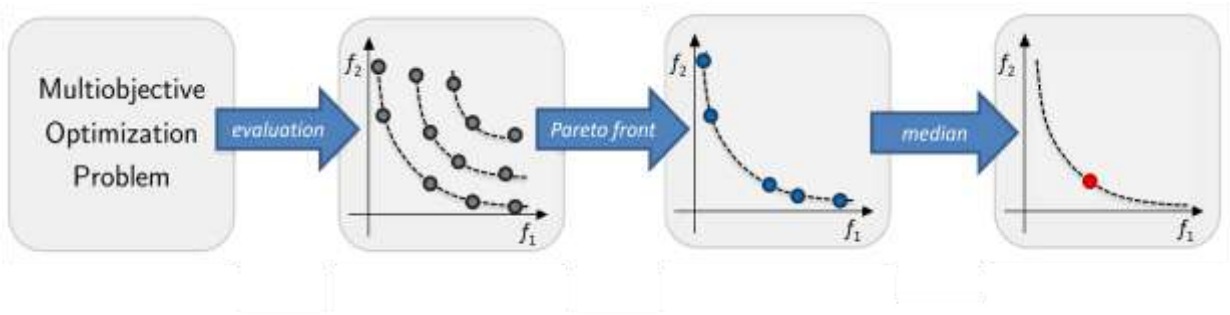


Figure 5.2. Illustration of multiobjective genetic procedure.

Concerning the first component, we will consider a population of M chromosomes $\alpha_m \in \mathbb{R}$, $m \in \{1, 2, \cdots, M\}$, where each chromosome is a real vector composed of genes corresponding to the weight vector $\boldsymbol{\alpha}$ defined above in the previous sections (see Figure 5.3). The length w of the chromosome is thus equal to the one of the dictionary D. Chromosomes can be randomly initialized or, to obtain a faster and better convergence, it could be envisioned to add *a priori* information coming from more simple CS techniques, i.e., OMP and BP algorithms. In other words, one could exploit OMP and BP solutions to generate an initial population by perturbing these solutions.

Regarding the fitness function, we will investigate separately and jointly two fitness functions, i.e., those defining the optimization problem in Eq. (5.3). The first one aims at maximizing the sparsity level by minimizing the L_0-norm of the weight vector α, which corresponds to minimizing the number of almost nonzero coefficients in α:

$$f_1 = min\|\boldsymbol{\alpha}\|_0 \,. \tag{5.7}$$

An almost nonzero coefficient is a coefficient exhibiting a value less than a predefined small threshold value (th). The second fitness function is derived from the constraint in Eq. (5.3). It

points to a perfect reconstruction of the considered pixel (at position $[k, l]$). In other words, it is expressed as:

$$f_2 = min\|D\boldsymbol{\alpha} - x\|^2. \quad (5.8)$$

NSGA-II returns several optimal (nondominated) solutions along the Pareto front. Since a single solution has to be selected from the nondominated set, different strategies can be adopted. In this study, we suggest to choose the median solution as typically performed in the literature (see last step in Figure 5.2). In such a way, we expect to get a compromise solution with respect to what could be obtained by OMP and BP, i.e., a tradeoff between reconstruction model sparseness and reconstruction error.

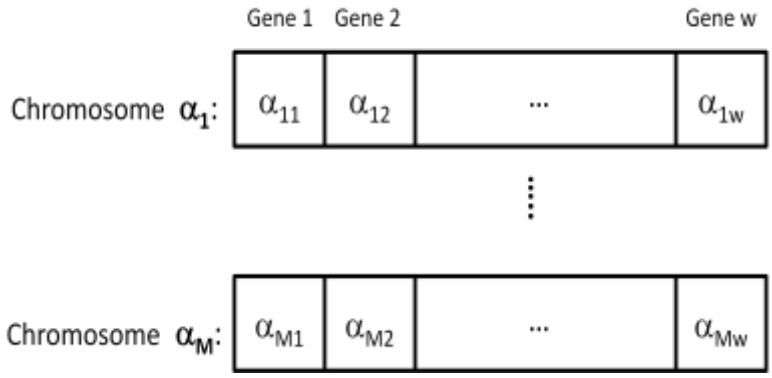


Figure 5.3. Adopted chromosome structure.

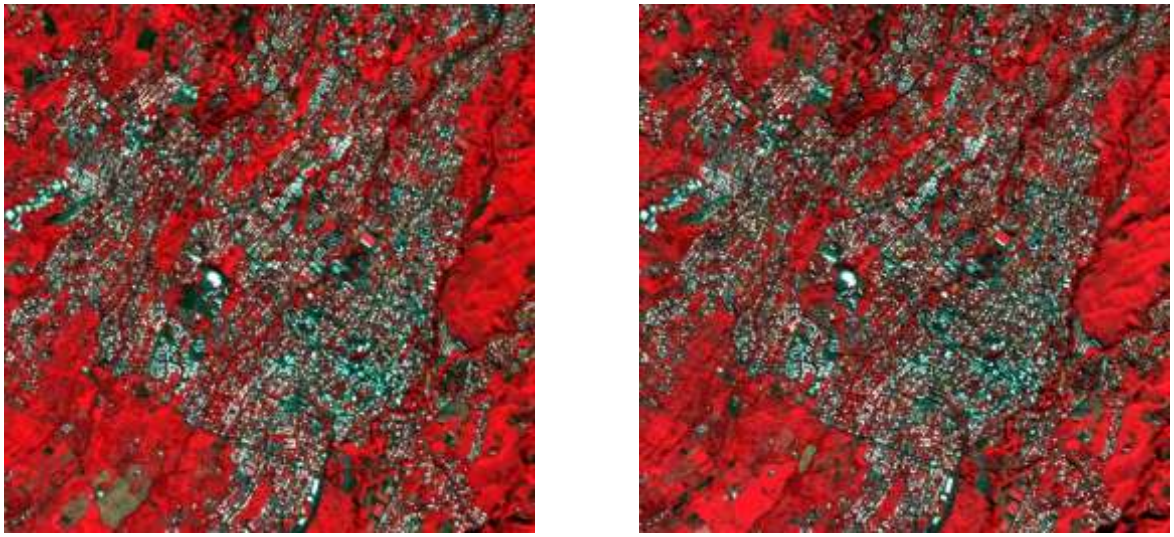
5.5. Experimental Results

5.5.1. Dataset description and setup

The first dataset comes from the Taiwanese optical high resolution FORMOSAT-2 satellite, which permits to acquire repeat imagery of an area of interest every day, from the same angle and under the same light conditions [39]. These images represent part of the *Arcachon* basin in the south region of Aquitaine, in France. The images are composed of 400×400 pixels, 4 spectral bands (blue, green, red and near-infrared) with a pixel spacing of 8 meters. They were acquired on the 24th of June and the 16th of July 2009, respectively (see Figure 5.4). The second dataset comes from the SPOT-5 French satellite, whose images represent part of the *Reunion* Island [40]. The images are characterized by a size of 450×450 pixels, 4 spectral bands (blue, green, red and near-infrared), a pixel spacing of 10 meters and were taken on the 2nd of May and the 18th of June 2008, respectively (see Figure 5.5). The two datasets present thus several differences among which the sensor type, pixel spacing, and the land covers. Indeed, the first one presents more vegetation areas than urban areas, while the second one exhibits an opposite composition. For the purpose of comparison, we implemented two other methods developed to reconstruct cloudy areas in RS images. One consists in a recent work, exploiting a multiresolution processing inpainting (MRI) [2], whereas the second method estimates a missing pixel by contextual multiple linear regression (CMLP) [5].

(a) (b)

Figure 5.4. Dataset 1: FORMOSAT-2 images acquired in the Arcachon area on (a) 24th of June and (b) the 16th of July, 2009.



(a) (b)

Figure 5.5. Dataset 2: SPOT-5 images acquired in the Reunion island on (a) the 2nd of May and (b) the 18th of June, 2008.

In order to make it possible to quantify the reconstruction accuracy of the methods, the experimental procedure adopted consisted: 1) to consider a cloud-free image, e.g., $I^{(1)}$; 2) to simulate the presence of clouds by partly obscuring the other image, e.g., $I^{(2)}$; and 3) to compare the reconstructed image with the original cloud-free image. The simulation study is aimed at understanding the sensitivity of the five investigated methods (i.e., the MRI, the CMLP, the OMP, the BP, and the GA methods) to two aspects, which are: 1) the kind of ground covers obscured; and 2) the size of the contaminated area. In order to obtain a detailed assessment of the reconstruction quality, we adopt the well-known peak signal-to-noise ratio (PSNR) measure [41]. Other quantitative criteria are the computation time (in seconds) and the model complexity, namely the number of coefficients required for the reconstruction model. Regarding the dictionaries, we collected directly training samples from the image, by sampling pixels in the source region $\mathbf{\Phi}$. For the GA setup, the parameters adopted by our experiment were as follows: chromosomes number

$M = 50$, threshold value $th = 10^{-4}$, probability of crossover $P_c = 0.8$ and probability of mutation $P_m = 0.005$.

(a) (b)

Figure 5.6. Masks adopted to simulate the contamination of different ground covers.

5.5.2. Results

1) Contamination of Different Ground Cover

Figure 5.6 shows different masks whose positions were selected in a way to simulate the obscuration of different kinds of ground cover. In particular, for dataset 1, Figure 5.6(a) shows mask A covering a region that includes mainly a urban area, mask B obscuring an industrial zone, and mask C covering a vegetation area. For dataset 2, Figure 5.6(b) shows mask A covering mainly a rural area, and mask B a vegetation region. The experiments were carried out by considering each mask at a time, where each mask is composed by around 2000 pixels, and the dictionary by 300 pixels.

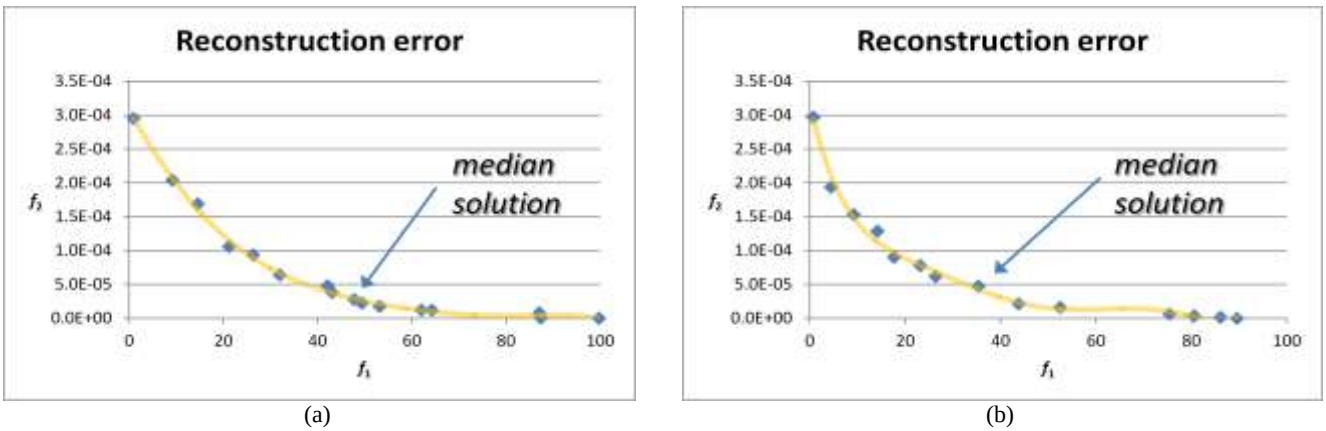
TABLE 5.I
QUANTITATIVE RESULTS OBTAINED IN THE FIRST SIMULATION EXPERIMENTS FOR DATASET 1.

Method	Mask A				Mask B				Mask C			
	PSNR		Complexity	time [s]	PSNR		Complexity	time [s]	PSNR		Complexity	time [s]
	I_1	I_2			I_1	I_2			I_1	I_2		
MRI	-	22.54	-	2856	-	16.05	-	2517	-	33.77	-	2898
CMLP	-	20.99	1	1	-	20.11	1	1	-	24.05	1	1
OMP	39.41	**23.96**	3	4	36.33	20.60	3	4	44.28	31.97	3	4
BP	80.59	22.22	294	66	77.10	**24.74**	168	59	98.53	30.67	301	60
GA	42.09	23.78	148	68621	43.38	23.15	95	26312	45.62	**32.01**	138	43193

TABLE 5.II
QUANTITATIVE RESULTS OBTAINED IN THE FIRST SIMULATION EXPERIMENTS FOR DATASET 2.

Method	Mask A				Mask B			
	PSNR		Complexity	time [s]	PSNR		Complexity	time [s]
	I_1	I_2			I_1	I_2		
MRI	-	24.27	-	2995	-	29.54	-	3614
CMLP	-	24.61	1	1	-	27.69	1	1
OMP	46.53	26.36	3	5	54.49	30.43	3	5
BP	86.22	26.45	338	61	99.62	**31.63**	365	91
GA	50.70	**26.72**	173	69231	56.30	31.28	201	38475

Table 5.I and Table 5.II report for the two datasets the results of the different reconstruction techniques over different obscured land covers. In greater detail, MRI generally reconstructs the missing data with a good PSNR level, but the corresponding reconstructed images appear visually of poor quality since it does not capture satisfactorily the textural properties of the missing areas. In general, MRI can return visually satisfactory results only when the missing area refers to a uniform region such as vegetation region. This is the case for mask C in dataset 1 and mask B in dataset 2. CMLP method provides generally satisfactory results in terms of reconstruction error and computation time. To obtain better results, it would need more than two temporal images. Coming now to our implementations, OMP algorithm produces very sparse reconstruction solutions (around 3 nonzero coefficients). On the one hand, this may be an advantage in terms of computation time. On the other hand, OMP is potentially subject to underfitting problems. On the contrary, BP algorithm may be subject to overfitting problems due to the fact that most of the time it selects a large number of weight coefficients (in general around 300 coefficients). Comparing OMP and BP in terms of computation time, the latter is far less efficient, whereas in terms of PSNR both methods return similar reconstruction values, outperforming CMLP and MRI. Lastly, GA can be viewed as a compromise between the two previous methods. Despite the very long time needed to estimate the reconstruction model, it results sparser than BP, but less parsimonious to OMP (see model complexity columns in Tables 5.I and 5.II). Its reconstruction error is almost all the time the best or the second best. Figure 5.7 illustrates the Pareto fronts obtained at convergence for mask A in both



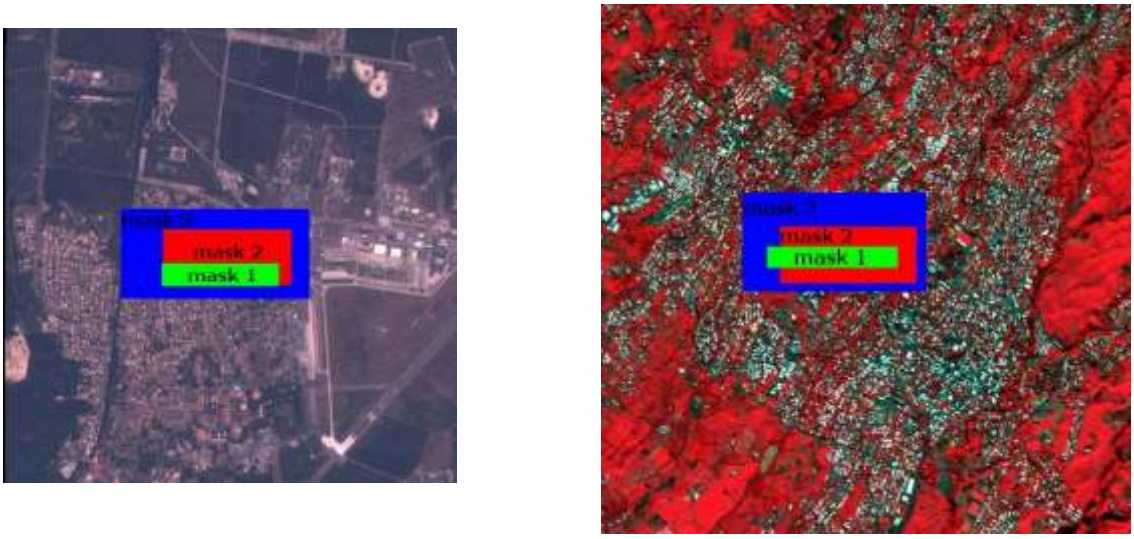
(a) (b)

Figure 5.7. Pareto fronts obtained at convergence for (a) dataset 1, and (b) dataset 2. Note that f_1 is represented in percentage.

datasets, and the corresponding median solutions. In general, it can be observed that satisfactory results were achieved with CS solutions despite the complexity of the faced problem.

2) Contamination With Different Sizes

Another important test for the five methods consists of assessing their performances by varying the amount of missing data. Figure 5.8 illustrates the three different masks adopted to simulate different increasing cloud cover sizes. In particular, masks 1 is the same as the masks A adopted in the previous experiments, i.e., it covers about 2000 pixels. To build masks 2 and 3, we multiplied the previous size, by 3 and by 6, obtaining masks covering around 6000 and 12000 pixels, respectively. Also in these experiments, the adopted dictionaries are composed of 300 pixels belonging to **Φ** region. Tables 5.III and 5.IV report for the two datasets the results achieved by the different reconstruction techniques and by varying the amount of missing data.



(a) (b)

Figure 5.8. Masks adopted to simulate the different sizes of contamination.

TABLE 5.III
QUANTITATIVE RESULTS OBTAINED IN THE SECOND SIMULATION EXPERIMENTS FOR DATASET 1.

Method	Mask 1				Mask 2				Mask 3			
	PSNR		Complexity	time [s]	PSNR		Complexity	time [s]	PSNR		Complexity	time [s]
	I_1	I_2			I_1	I_2			I_1	I_2		
MRI	-	22.54	-	2856	-	21.35	-	6938	-	19.63	-	14774
CMLP	-	20.99	1	1	-	21.13	1	1	-	20.83	1	2
OMP	39.41	**23.96**	3	4	42.45	23.21	7	6	42.00	**25.01**	3	19
BP	80.59	22.22	294	66	80.18	22.89	277	145	79.53	21.47	265	865
GA	42.09	23.78	148	68621	45.46	**23.85**	140	99072	45.13	23.03	149	275394

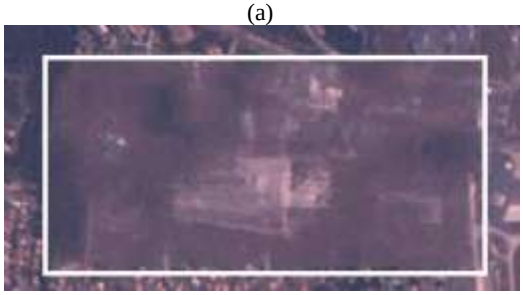
TABLE 5.IV
QUANTITATIVE RESULTS OBTAINED IN THE SECOND SIMULATION EXPERIMENTS FOR DATASET 2.

Method	Mask 1				Mask 2				Mask 3			
	PSNR		Complexity	time [s]	PSNR		Complexity	time [s]	PSNR		Complexity	time [s]
	I_1	I_2			I_1	I_2			I_1	I_2		
MRI	-	24.27	-	2995	-	22.85	-	10176	-	23.82	-	22353
CMLP	-	24.61	1	1	-	24.43	1	2	-	25.46	1	2
OMP	46.53	26.36	3	5	46.89	26.42	3	16	47.49	27.39	3	21
BP	86.22	26.45	338	61	87.49	26.82	332	143	86.60	**28.25**	329	972
GA	50.70	**26.72**	173	69231	50.63	**27.10**	168	103342	51.14	28.15	170	259459

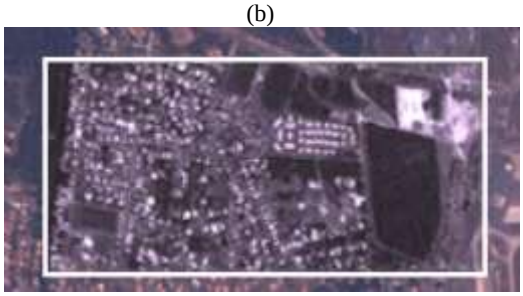
From a quantitative viewpoint, in terms of PSNR, we have similar results as in the previous experiments. MRI still presents problems in reconstructing satisfactorily complex textures. CMLP competes seriously with MRI in terms of computation time and PSNR. However to get higher PSNR values, one needs to resort to CS techniques. Indeed, our implementations return better results in term of PSNR in all the simulations and present the advantage for not depending on the size of the missing area. The best solution in these experiments in terms of PSNR comes from GA, which outperforms all other methods in three cases, and in the other three it is the second best choice. About the computation time, as expected, it increases as the amount of missing data increases. Results from this viewpoint underline the main weakness of the GA solution, i.e., its expensive computational needs. Regarding the model complexity, we got in these experiments results in line with those of the previous series of experiments.

Figures 5.9 and 5.10 show qualitative reconstruction results in RGB composites obtained in the most critical reconstruction scenario, i.e., the largest simulated cloud mask 3, for both datasets and for all reconstruction methods. As mentioned before, MRI reconstruction exhibits the worst reconstruction case. CMLP method is capable to obtain a good reconstruction compared with MRI. Regarding the CS reconstruction techniques (OMP, BP, and GA), good reconstructions are obtained, especially for dataset 2, where it is not simple to find significant differences comparing the reconstructions with the original (cloud-free) images.

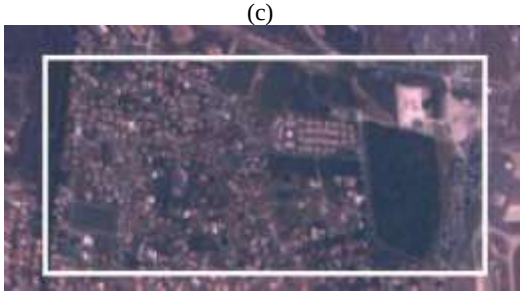
(a)



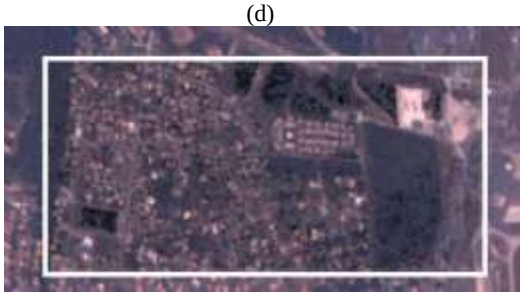
(b)



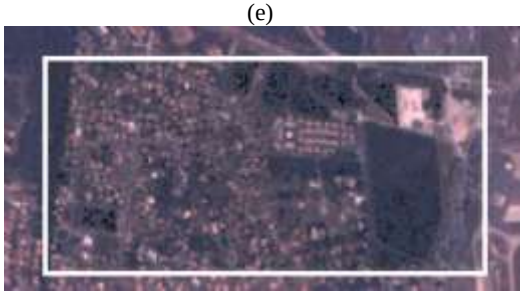
(c)



(d)

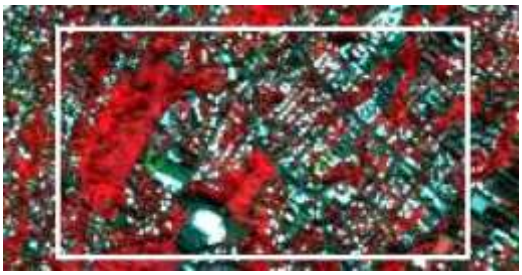


(e)

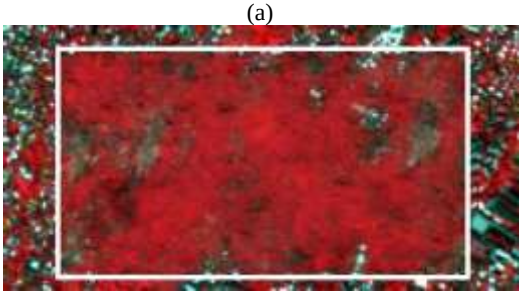


(f)

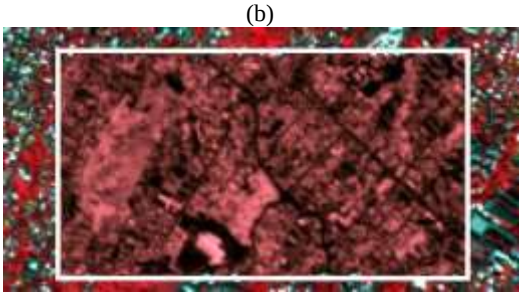
Figure 5.9. Dataset 1. Color composite images (bands 1, 2, and 3) (a) of the original image, and the same image reconstructed after the contamination with the largest simulated mask 3 by (b) MRI, (c) CMLP, (d) OMP, (e) BP, and (f) GA methods.



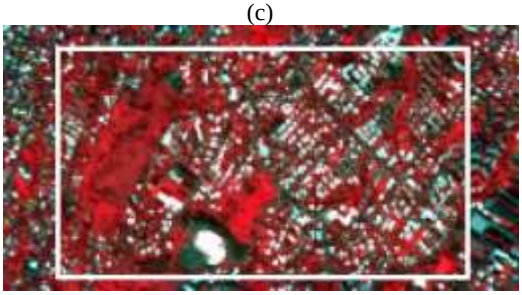
(a)



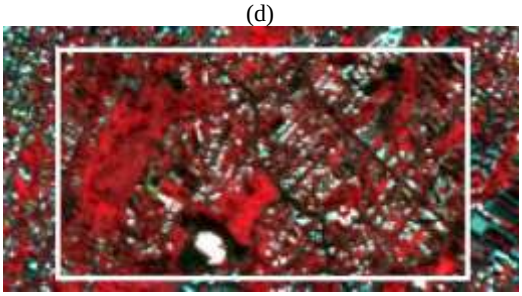
(b)



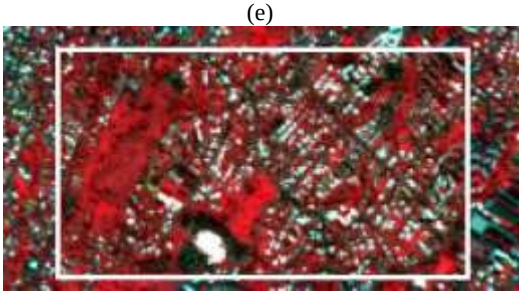
(c)



(d)



(e)



(f).

Figure 5.10. Dataset 2. Color composite images (bands 1, 2, and 3) (a) of the original image, and the same image reconstructed after the contamination with the largest simulated mask 3 by (b) MRI, (c) CMLP, (d) OMP, (e) BP, and (f) GA methods.

5.5.3. Reconstruction impact on image classification

Since the generation of classification maps represents one of the most widespread applications of remote sensing images, it was also worth to evaluate the quality of the reconstruction process in terms of classification error. The latter was computed first by generating a classification map of the original images that served as reference classification maps by means of the popular *k*-means classifier. Then, each reconstructed image was given in input to the *k*-means classifier to provide a reconstruction classification map. For each reconstruction method, it was thus possible to evaluate the overall number of classification errors (OE) inside the reconstructed cloud-contaminated area by a simple comparison of both the reconstruction and the original classification maps. We repeated this exercise with different numbers of clusters (from 3 to 7 clusters). The results confirm what previously observed, i.e., compressive sensing methods behave better than traditional ones. As example, we have reported in Figure 5.11 the clustering results (with k=5) obtained for the FORMOSAT-2 image with mask A. Figure 5.12 shows in more details the differences between the various methods at the level of the reconstructed area (i.e., the mask). The best classification is achieved from the reconstruction with OMP (OE of 6.3%), followed by GA (OE=8.7%), BP (OE=19.6%), MRI (OE=25.2%) and CMLP (OE=29.4%).

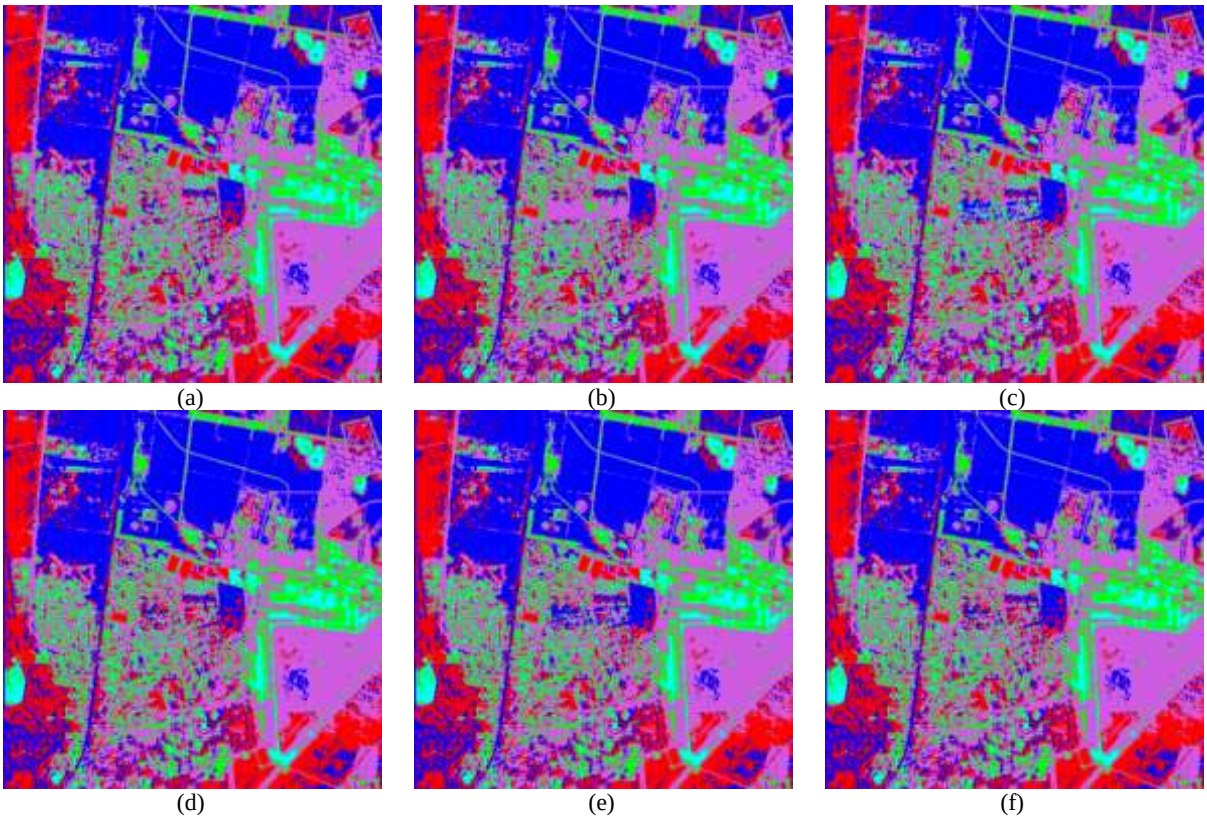


Figure 5.11.Unsupervised classification maps obtained by the k-means algorithm (k=5) (a) for the original FORMOSAT-2 image; and the same image reconstructed after contamination with mask A by (b) MRI, (c) CMLP, (d) OMP, (e) BP, and (f) GA methods.

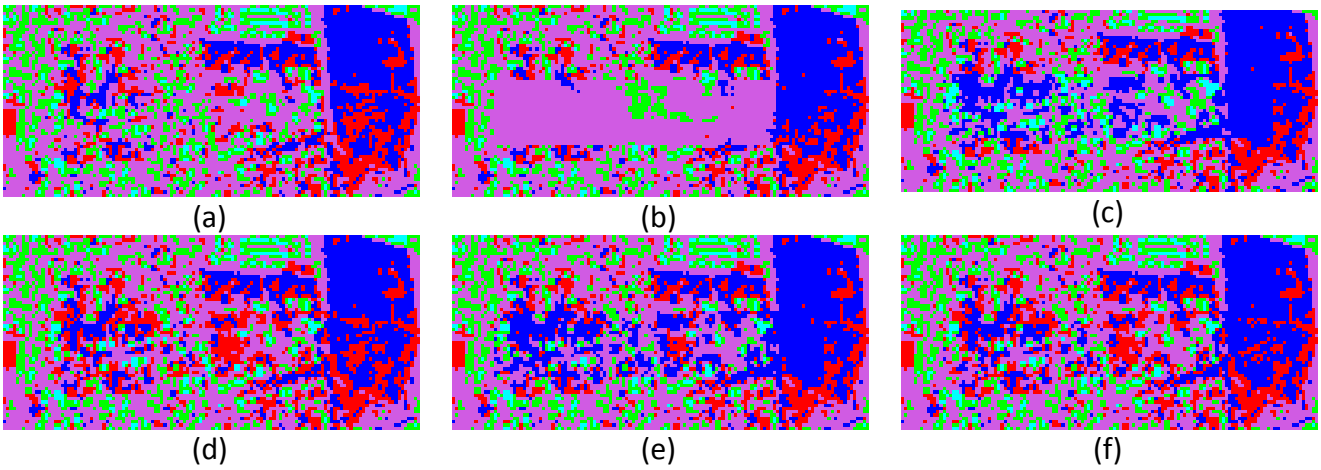


Figure 5.12. Zooming of Figure 5.11 on mask A area: (a) original classification and classification after reconstruction with (b) MRI, (c) CMLP, (d) OMP, (e) BP, and (f) GA methods.

5.6. Conclusion

This paper deals with the complex and important problem of the removal of clouds from multispectral images. In particular, three novel methods have been proposed, which, given a contaminated image and a cloud-free image, allow reconstructing the missing measurements through the compressive sensing approach. First, we have shown how two common CS solutions, the orthogonal matching pursuit (OMP) and the basis pursuit (BP) algorithms, can be formulated for a cloud-contaminated image reconstruction problem. Then, we have proposed a solution for solving the CS problem under a L_0–norm perspective, exploiting the capabilities of genetic algorithms. The main properties of the proposed methods are: 1) they rely on the assumptions that spectral nonstationarity is allowed, while the spatial structure of the image should be almost identical between the two images; 2) they are not ground cover-dependent; 3) they are unsupervised; 4) differently from CMLP, they need just one reference cloud-free image and the reconstruction of each pixel is performed in all spectral bands simultaneously.

The experimental results point out the superiority of the proposed methods compared to two reference methods for cloud removal. Comparing our proposed methods, all three exhibit good results in the reconstruction of missing areas. OMP has the advantage to be sparser and significantly faster than BP and GA, but it is the less robust method. Indeed, since the reconstruction of each pixel depends typically on 3 coefficients and thus 3 other pixels of the image, it is enough that one of them is missing (covered by a cloud) to render the reconstruction model inaccurate. This problem is much less important to BP as it is much less sparse than OMP. GA represents a good compromise between OMP and BP methods, mainly because it is more robust than OMP and more sparse than BP. Another empirical conclusion is that the kind of ground cover obscured may be an important factor to take in consideration for the reconstruction, while the size of the contaminated area only marginally affects the performance of the proposed reconstruction methods, which depends more on the information available outside the missing area. In other words, if a ground cover is contaminated and it is not represented outside of the contaminated area, the reconstruction process will not deal suitable with such a situation.

Finally, in order to improve the accuracy of the reconstruction process, different aspects of the methods deserve to be investigated in future research studies. For example, in almost all the reconstructed images we note an unwanted presence of a slight "curtain effect" because of the presence of an estimation bias. A possible solution to this problem could be a cluster-oriented sparse reconstruction. Note that, in general, it is difficult to obtain pairs of images where one of them is completely cloud-free. Adopting the implemented CS solutions, it will be possible to exploit and reconstruct two cloud-contaminated images only if they do not convey clouds in the same region.

5.7. Acknowledgement

The authors are grateful to CNES for making available on line the FORMOSAT-2 and SPOT-5 images (http://kalideos.cnes.fr). The authors would like to thank D. Donoho and Y. Tsaig for the SparseLab toolbox used in this research (http://sparselab.stanford.edu/).

5.8. References

[1] S. C. Liew, M. Li, and L. K. Kwoh, "Automated Production of Cloud-free and Cloud-shadow Image Mosaics from Cloudy Satellite Imagery", in Proc. XXth ISPRS Congr., Istanbul, Turkey, Jul. 12–13, pp. 523-530, 2004.

[2] L. Lorenzi, F. Melgani and G. Mercier, "Inpainting Strategies for Reconstruction of Missing Data in VHR Images", IEEE Geosci. Remote Sens. Letters, vol. 8, no. 5, pp. 914-918, Sep. 2011.

[3] A. Maalouf, P. Carre, B. Augereau, and C. Fernandez-Maloigne, "A bandelet-based inpainting technique for clouds removal from remotely sensed images," IEEE Trans. Geosci. Remote Sens., vol. 47, no. 7, pp. 2363–2371, Jul. 2009.

[4] S. Lee and M. M. Crawford, "An Adaptive Reconstruction System for Spatially Correlated Multispectral Multitemporal Images", IEEE Trans. Geosci. Remote Sens., vol. 29, no. 4, pp. 494–508, Jul. 1991.

[5] F. Melgani, "Contextual Reconstruction of Cloud-Contaminated Multitemporal Multispectral Images", IEEE Trans. Geosci. Remote Sens., vol. 44, no. 2, pp. 442–455, Feb. 2006.

[6] D-C. Tseng, H-T. Tseng and C-L. Chien, "Automatic Cloud Removal From Multi-Temporal SPOT Images", Appl. Math. Comput., vol. 205, no. 2, pp. 584-600, Nov. 2008.

[7] C-H. Lin, P-H. Tsai, K-H. Lai and J-Y. Chen, "Cloud Removal From Multitemporal Satellite Images Using Information Cloning", IEEE Trans. Geosci. Remote Sens., vol. 50, in press, 2012.

[8] D. L. Donoho, "Compressed Sensing", IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[9] E. J. Candès, J. Romberg, and T. Tao, "Robust Uncertainty Principle-Exact Signal Reconstruction from Highly Incomplete Frequency Information", IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[10] L. Jiying, Z. Jubo, Y. Fengxia and Z. Zenghui, "Theoretical Frameworks of Remote Sensing Systems Based on Compressive Sensing", International Archives of Photogrammetry and Remote Sensing, vol. 37, pp. 77-81, Jul. 2010.

[11] M. Aharon, M. Elad and A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation", IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4311-4322, Nov. 2006.

[12] J. Wright, A. Yang, A. Ganesh, S. Sastry and Y. MA, "Robust Face Recognition via Sparse Representation", IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 2, pp. 210-227, Feb. 2009.

[13] J. Yang, J. Wright, T. Huang and Y. Ma, "Image Super-Resolution Via Sparse Representations", IEEE Trans. on Image Process., vol. 19, no. 11, pp. 2861-2873, Nov. 2010.

[14] S. Rao, R. Tron, R. Vidal and Y. Ma, "Motion Segmentation via Robust Subspace Separation in the Presence of Outlying, Incomplete, and Corrupted Trajectories", IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 10, pp. 1832-1845, 2008.

[15] J. Mairal, M. Elad and G. Sapiro, "Sparse Representation for Color Image Restoration", IEEE Trans. On Image Process., vol. 17, no. 1, pp. 53-69, Jan. 2008.

[16] B. Shen, W. Hu, Y. Zhang and Y.-J. Zhang, "Image Inpainting via Sparse Representation", in IEEE ICASSP, pp. 697-700, 2009.

[17] A. Quattoni, M. Collins and T. Darrell, "Transfer Learning for Image Classification with Sparse Prototype Representation", In Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., pp. 1-8, Jun. 2008.

[18] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T.S. Huang, and S. Yan, "Sparse Representation for Computer Vision and Pattern Recognition", in Proc. of the IEEE, vol. 98, no. 6, Jun. 2010.

[19] S. Li and B. Yang, "A New Pan-Sharpening Method Using a Compressed Sensing Technique", IEEE Trans. Geosci. Remote Sens., vol. 49, no. 2, pp. 738-746, Feb. 2011.

[20] Y. Chen, N. M Nasrabadi and T. D. Tran, "Hyperspectral Image Classification Using Dictionary-Based Sparse Representation", IEEE Trans. Geosci. Remote Sens., vol. 39, no. 10, pp. 3973-3985, Oct. 2011.

[21] S. Yang, F. Sun, M. Wang, Z. Liu and L. Jiao, "Novel Super Resolution Restoration of Remote Sensing Images Based on Compressive Sensing and Example Patches-aided Dictionary Learning", International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping, pp. 1-6, Jan 2011.

[22] J. Ma, "Single-Pixel Remote Sensing", Geosci. Remote Sens. Letters, vol. 6, no. 2, pp. 199-203, Apr. 2009.

[23] E. J. Candès and T. Tao, "Decoding by Linear Programming", IEEE Trans. Inform. Theory, vol. 51, no. 12, pp. 4203-4215, Dec. 2005.

[24] S. S. Chen, D. L. Donoho and M. A. Saunders, "Atomic Decomposition by Basis Pursuit", SIAM J. on Sci. Comp., vol. 20, pp. 33-61, 1999.

[25] A. M. Bruckstein, D. L. Donoho and M. Elad, "From Sparse Solutions of Systems of Equations to Sparse Modelling of Signals and Images", SIAM, vol. 51, pp. 34-81, Feb. 2009.

[26] N. Wang and Y. Wang, "An Image Reconstruction Algorithm based on Compressive Sensing Using Conjugate Gradient", in Proc. IUCS 2010, pp. 374-377, Oct. 2010.

[27] Y. C. Pati, R. Rezaiifar and P. S. Krishnaprasad, "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decompositions", in Proc. 27th Asilomar Conf. on Sig., Sys. and Comp., 1993.

[28] J. A. Tropp and A. C. Gilbert, "Signal Recovery from Random Measurements via Orthogonal Matching Pursuit", IEEE Trans. Inform. Theory, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.

[29] S. Kunis, and H. Rauhut, "Random Sampling of Sparse Trigonometric Polynomials, II. Orthogonal Matching Pursuit versus Basis Pursuit", Found. of Comput. Math., vol. 8, no. 6, pp. 737-763, Dec. 2008.

[30] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning. Reading, MA: Addison-Wesley, 1989.

[31] L. Chambers, The Practical Handbook of Genetic Algorithms. New York: Champan & Hall, 2001.

[32] K. Deb, Multi-Objective Optimization Using Evolutionary Algorithms. Chichester, U.K.: Wiley, 2001.

[33] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength pareto evolutionary algorithm," in Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems, pp. 95–100, 2002.

[34] J. Knowles and D. Corne, "The Pareto archived evolution strategy: A new baseline algorithm for pareto multiobjective optimisation," in Proceedings of the Congress on Evolutionary Computation (CEC'99), vol. 1, pp. 98–105, 1999.

[35] N. Srinivas and K. Deb, "Multiobjective function optimization using nondominated sorting genetic algorithms," Evol. Comput., vol. 2, no. 3, pp. 221–248, Fall 1995.

[36] N. Ghoggali, F. Melgani, and Y. Bazi, "A Multiobjective Genetic SVM Approach for Classification Problems With Limited Training Samples", IEEE Trans. Geosci. Remote Sens., vol. 47, no. 6, pp. 1707–1718, Jun. 2009.

[37] N. Ghoggali, and F. Melgani, "Genetic SVM Approach to Semisupervised Multitemporal Classification", IEEE Geosci. Remote Sens. Letters, vol. 5, no. 2, pp. 212-216, Apr. 2008.

[38] E. Pasolli, F. Melgani, and M. Donelli, "Automatic Analysis of GPR Images: A Pattern-Recognition Approach", IEEE Trans. Geosci. Remote Sens., vol. 47, no. 7, pp. 2206–2217, Jul. 2009.

[39] C-C. Liu, "Processing of FORMOSAT-2 Daily Revisit Imagery for Site Surveillance", IEEE Trans. Geosci. Remote Sens., vol. 44, no. 11, pp. 3206–3215, Nov. 2006.

[40] A. Baudoin, "Mission Analysis for SPOT 5", in IEEE IGARSS, vol. 3, pp. 1084, Aug. 1993.

[41] A. K. Jain, Fundamentals of Digital Image Processing. New York: Prentice Hall, 1988.

6. A Complete Processing Chain for Shadow Detection and Reconstruction in VHR Images

Abstract *– The presence of shadows in very high resolution (VHR) images can represent a serious obstacle for their full exploitation. This chapter proposes to face this problem as a whole through the proposal of a complete processing chain, which relies on various advanced image processing and pattern recognition tools. The first key point of the chain is that shadow areas are not only detected but also classified so that to allow their customized compensation. The detection and classification tasks are implemented by means of the state-of-the-art support vector machine (SVM) approach. A quality check mechanism is integrated in order to reduce subsequent misreconstruction problems. The reconstruction is based on a linear regression method to compensate shadow regions by adjusting the intensities of the shaded pixels according to the statistical characteristics of the corresponding non-shadow regions. Moreover, borders are explicitly handled by making use of adaptive morphological filters and linear interpolation for the prevention of possible border artifacts in the reconstructed image. Experimental results obtained on three VHR images representing different shadow conditions are reported, discussed and compared with two other reconstruction techniques.*

The work presented in this chapter has been has been published in the *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3440–3452, September 2012; Co-authors: F. Melgani and G. Mercier.

6.1. Introduction

Recently, very high resolution (VHR) satellite images opened a new era in the remote sensing field. Because of the increase of spatial resolution, new analysis, classification and change detection techniques are required. Indeed, VHR images exhibit resolutions which allow distinguishing very well detailed features from small objects, like little building structures, trees, vehicles and roofs. Unfortunately, high spatial resolution entails also some drawbacks like the unsought presence of shadows, particularly in urban areas where there are larger changes in surface elevation (due to the presence of buildings, bridges, towers, etc) and consequently longer shadows. Although it is feasible to exploit shadow characteristics to recognize building position, to estimate their height and other useful parameters ([1]-[2]), usually shadows are viewed as undesired information that strongly affects images. Shadows may cause a high risk to present false color tones, to distort the shape of objects, to merge or to lose objects. They represent an important problem for both, users and sellers of remote sensing images [2]. As a consequence, shadows can impact negatively in the exploitation of VHR images, influencing detailed mapping, leading to erroneous classification or interpretation (e.g., biophysical parameters such as vegetation, water or soil indexes), due to the partial or total loss of information in the image [3]. To attenuate these drawbacks and thus to increase image exploitability, two steps are necessary: 1) shadow detection; and 2) shadow compensation (reconstruction). An example of the importance of getting shadow-free images is the massive tsunami in 2004 where it was crucial to obtain such images in a very short time in order to take rapid and crucial decisions in rescue missions [4].

The literature reports mainly two approaches to detect shadows, namely model-based and shadow properties-based approaches. The former needs prior information about the scenario and the sensor. However, since usually such knowledge is not available, most of the detection algorithms are based on shadow properties, such as the fact that shadow areas have lower brightness, higher saturation and greater hue values [5]. For instance, methods in [6] and [7] attempt to detect shadows using a space color transformation and an automatic threshold estimator (e.g., Otsu's algorithm [8]). In a comparative work [9], several invariant color spaces were analyzed to detect shadows, namely HIS, HSV, HCV, YIQ and YCbCr transforms. Inspired by this comparative analysis, a better approach was developed, which is based on a novel successive thresholding scheme [10]. Other algorithms rely on the idea of adding features capable to better discriminate shadow areas (e.g., NDVI normalized difference vegetation index [11], NSVDI normalized saturation-value difference index [12], MSER maximally stable extremal regions [13]). Another technique applies the principal component analysis (PCA) to isolate the luminance component in an RGB image, where the detection of shadows appears more accurate [14]. Finally, physical properties (e.g., temperature) of a blackbody radiator have been exploited in a recent method to detect shadows [13].

In order to compensate/reconstruct shadow areas, there exist essentially three different methods: 1) gamma correction; 2) histogram matching; and 3) linear correlation [16]. In [17], the authors assume that the surface texture does not radically change when it is shaded. Accordingly, to remove shadows, they perform a contextual texture analysis between a segment of shadow and its neighbors. Knowing the kind of surface under the shadow, a local gamma transformation is then used to restore the shadow area. In [18], after the detection of shadows, the authors propose to

adjust the hue, intensity and saturation values (HIS) in shadow regions respectively according to the analogous values in the local surrounding of each shadow region, by adopting the histogram matching method. In [19], the method consists to recover spectral information in shadow areas in an IKONOS image by exploiting the height data from the airborne laser scanner (ALS). Such information is used to overlay and eliminate the real shadow. The results obtained with two methods, namely gamma correction and linear correlation, are compared. In [20], it is assumed that the restoration of shadows almost depends on the spectral signature of the spectral bands. Accordingly, first the bands are thresholded in an independent way determining the optimal threshold values by visual inspection. Then, a linear regression in each spectral band is carried out to correct the shadow effects.

In this chapter, an alternative method is proposed to solve both problems of detection and reconstruction of shadow areas. Shadow detection is performed through a hierarchical supervised classification scheme, while the proposed reconstruction relies on a linear correlation function, which exploits the information returned by the classification. The whole processing chain includes also two important capabilities: 1) a rejection mechanism to limit as much as possible reconstruction errors; and 2) explicit handling of the shadow borders.

The remainder of this chapter is organized as follows. In Section 6.2, the problem of the presence of shadows in VHR images is formulated. Section 6.3 details our approach. Section 6.4 shows experimental results and Section 6.5 draws the conclusions.

6.2. Problem Formulation

In VHR optical images, especially in urban areas, the presence of shadows may completely destroy the information contained in those images. Information missing in shadow areas directly influences common processing and analysis operations, such as the generation of classification maps [21]-[22]. Normally shadows appear when objects occlude the direct light from the illumination source, usually the sun. But shadows are not all the same, they can be divided in two different classes: cast and self shadows (see Figure 6.1). *Cast shadow* (following the terminologies used in [13], [15]) is caused by the projection of the light source in the direction of the object. *Self-shadow* is still a shadow but represents the part of the object that is not illuminated directly by the light source [13], [15]. It can come from the diffuse light present in the scene, and it may have a nonlinear behaviour. For simplicity, this work does not distinguish between self and cast shadows. It assumes that most of the shadows in a given image belong to the cast type, producing homogeneous dark areas with a loss of information that we desire to recover.

In this chapter, the detection of shadows is made through a hierarchical supervised classification process in order to: 1) first separate between shadow and non-shadow areas in the given image; and then 2) identify the different non-shadow classes as well as their corresponding shadow counterpart. Accordingly, ground truth is needed for both categories of classes. Note that it is the only part where the human help is needed. The reconstruction of shadow areas is based on the hypothesis that both shadow and non-shadow classes follow a Gaussian distribution. Though it can be expected that such a hypothesis does not always hold, it is however useful so that to get a simple

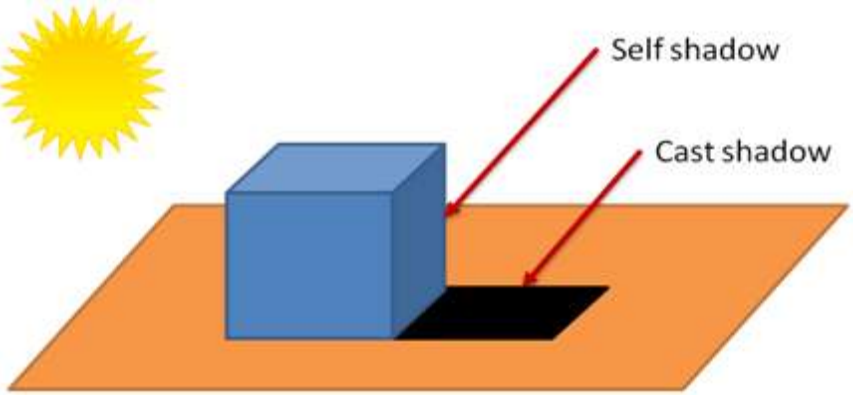


Figure 6.1. Illustration of cast and self shadows

and fast solution to the reconstruction problem. Indeed, denoting the shadow class as $X \sim \mathrm{N}\left(\mu_{\mathrm{S}}, \sigma_{\mathrm{S}}^2\right)$ and the corresponding non-shadow class as $Y \sim \mathrm{N}\left(\mu_{\bar{\mathrm{S}}}, \sigma_{\bar{\mathrm{S}}}^2\right)$, the reconstruction of the shadow class will be reduced to a simple random variable transformation:

$$X \sim \mathrm{N}\left(\mu_{\mathrm{S}}, \sigma_{\mathrm{S}}^2\right) \rightarrow X' \sim \mathrm{N}\left(\mu_{\bar{\mathrm{S}}}, \sigma_{\bar{\mathrm{S}}}^2\right) . \tag{6.1}$$

6.3. Proposed Method

Figure 6.2 depicts a flow chart with the principal steps of the proposed methodology. In brief, let us consider a VHR image *I* of dimensions $m \times n$, composed of *N* bands and characterized by the presence of shadow areas. As a first step, ground-truth information is collected by selecting different regions of interest (ROI) in order to discriminate the ("clean", non-shadow) classes present in the image as well as their shadow counterpart. It is noteworthy that this initial step is the only one where human interaction is required for building ROIs for all classes (shadowed or not). The resulting ground truth will allow performing first a binary classification in order to distinguish between shadow and non-shadow regions. To deal with noise which may result in the obtained binary mask *M*, two mathematical morphological operators are applied, namely opening and closing by reconstruction [23]-[24]. Because the binary mask does not handle the presence of the penumbra, and also may not return precise edges between the two classes, a border *B* between them is created, still by making use of morphological operators. The border areas are exploited in the last step of the processing chain for interpolation purposes. In a successive step, shadow and non-shadow classes are classified separately with the same initial ROI. Such classification allows the localization of the available couples of shadow and non-shadow related to the same object and thus to define the spectral relationship between them as a means to perform the reconstruction of the shadow areas. In particular, the reconstruction is based on a linear regression method to compensate shadow regions where the intensities of the shaded pixels are adjusted according to the statistical characteristics of the corresponding non-shadow regions. Finally, the border between the reconstructed shadow and the non-shadow areas undergoes a linear interpolation operation to yield a smooth transition between them. In the following paragraphs, a detailed description of all these steps is provided.

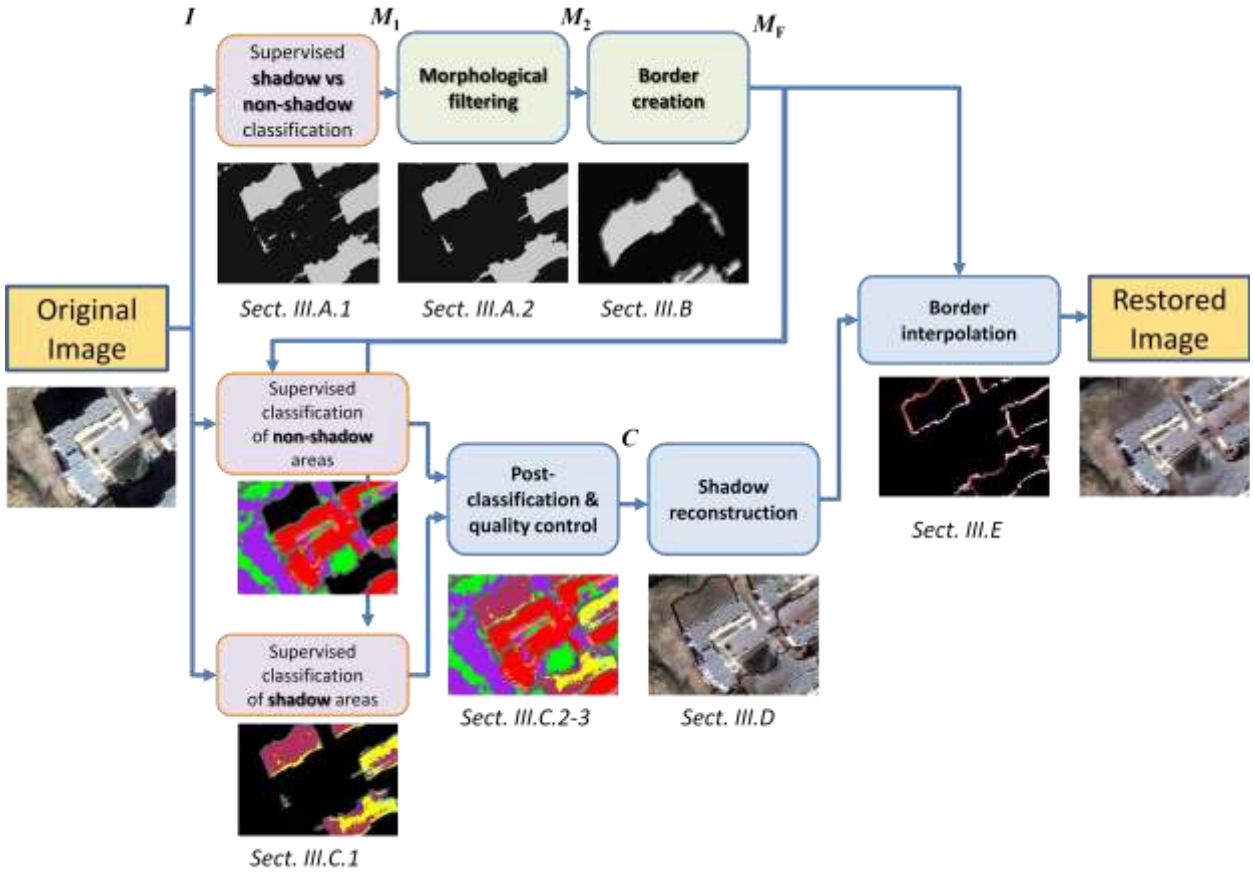


Figure 6.2. Flow chart of the proposed method.

6.3.1. Mask construction

The shadow *vs.* non-shadow mask is created in two steps, namely binary classification followed by a post-processing.

6.3.1.1. Binary classification

The binary classification procedure (see M_1 in Figure 6.3(a)) is implemented in a supervised way by means of a support vector machine (SVM), which proved its effectiveness in the literature of remote sensing data classification [25]-[27]. The feature space where to perform the classification task is defined by the original image bands and features extracted by means of the wavelet transform. In particular, a one-level stationary wavelet transform is applied on each spectral band, obtaining thus for each band four space-frequency features. The symlet wavelet is adopted in order to maximize the sparseness of the transformation (most of the coefficients are near 0) while enforcing texture areas (wavelet coefficients are of high value on presence of singularities) [28]. For an original image I composed of B spectral bands, the resulting feature space consists thus of $B\times(1+4)$ dimensions.

(a) (b) (c)

Figure 6.3. Example of (a) initial mask image M_1, (b) mask post-processed by opening and closing M_2, (c) final mask with borders M_F.

6.3.1.2. Post-processing

The binary image M_1 may be characterized by a "salt and pepper" effect due to the presence of noise in the image. An opening by reconstruction followed by a closing by reconstruction is applied on M_1 to attenuate this potential problem [23]. The choice of morphological filters to deal with this problem is motivated by their effectiveness and better shape preservation capability as shown in the literature, and by the possibility to adapt them according to the image filtering requirements as is the case in the border creation (described in the next subsection) [29]-[30]. Both morphological operators are needed in order to remove isolated shadow pixels in a non-shadow area and also isolated non-shadow pixels in a shadow area. For illustration, an example of enhanced mask (binary classification) M_2 is given in Figure 6.3(b), where a morphological structuring element (SE) of 3×3 pixel size is considered as shown in Figure 6.4(a).

6.3.2. Border creation

The transition in-between shadow and non-shadow areas can raise problems such as boundary ambiguity, color inconstancy, and illumination variation [9]. Indeed, the presence of the penumbra induces mixed pixels which are difficult to classify. The penumbra is a region where the light source is only partially obscured. For this reason, a border between the shadow and non-shadow classes is defined in order to appropriately handle the border pixels. These lasts are not processed within the shadow reconstruction procedure as is, but separately. The border region is constructed by means of morphological operators. The mask c_imgB2 is dilated (δ) and eroded (ε). Then, the difference between these two images is computed to form the border image *B*:

$$\mathrm{B[x,y]} = \delta(\mathrm{c\_imgB2[x,y]}) - \varepsilon(\mathrm{c\_imgB2[x,y]}). \quad (6.2)$$

Note that the border is not needed in all directions. Indeed, in order to only focus on the penumbra, the sun direction is taken into account to adapt a proper structure for the morphological operators so that to better track the actual shadow direction (see Figure 6.4(b)). The final mask image becomes (see example in Figure 6.3(c)):

$$c\_imgB_{NEW}[x,y] = \begin{cases} B[x,y] if B[x,y] = 1 \\ c\_imgB2[x,y] if B[x,y] = 0 \end{cases}. \quad (6.3)$$

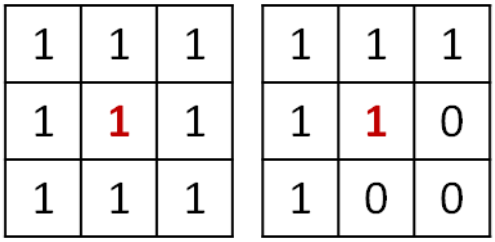


Figure 6.4. Examples of structuring elements used for (a) the standard morphological operator and (b) the morphological operator adapted to the shadow geometry related to the image in Figure 6.3.

6.3.3. Classification maps

6.3.3.1. Multiclass classification

The previously obtained mask is exploited to guide a further level of classification applied separately to the shadow and non-shadow areas. The aim is to distinguish between the different predefined non-shadow classes on the one side and the corresponding shadow classes on the other side. The result is a final classification map *C*, important to define the spectral relationship between the shadow and non-shadow versions of the same object (class) and thus to perform customized reconstruction of shadow areas. For such purpose, two multiclass SVMs are trained in the feature space described above for the shadow and non-shadow classifications, respectively. After the training phase, to generate *C*, both are applied to predict the label of each pixel of the corresponding areas, shadow and non-shadows respectively, defined in M_{F}.

6.3.3.2. Post-classification

In order to improve the classification map *C* before exploiting it for the reconstruction of shadow areas, post-classification is applied by adopting a simple 3×3 majority filter for removing isolated labels and thus smoothing the map.

6.3.3.3. Quality control

It can reasonably be expected that the classification of shadow areas is tricky because the spectral signatures of shadow classes have usually low and close radiometric behaviors. Since the reconstruction process is directly based on the classification map, it becomes necessary to control its quality in order to decide if compensation is feasible or not for a particular shadow class. For such purpose, a confusion matrix for the shadow classes is computed on the basis of the available ground-truth. For each shadow class, both user (corresponding to commission errors, or inclusion) and producer (corresponding to omission errors, or exclusion) accuracies are derived [31]. If one of these accuracies is lower than a predefined value, the classification of the corresponding class is

considered of low quality and, therefore, the related shadow compensation is not performed. In other words, shadow reconstruction is carried out only wherever a sufficient guaranty of correct shadow recognition is available so that to avoid error propagation in the processing chain.

6.3.4. Shadow reconstruction

Image reconstruction is one of the most important steps in our methodology. As done in the literature [14], [16], [19], [20], for the sake of getting a simple but satisfactory reconstruction model, we assume that the underlying relationship between the non-shadow class (Y) and the corresponding shadow classes (X) is of the linear type. We have empirically observed that shadow classes and the corresponding non-shadow classes reasonably exhibit a linear relationship. Regarding the statistical model of the classes, three estimation ways may be envisioned: 1) histogram estimation by box counting, 2) kernel density estimation (KDE) or 3) parametric estimation. In our case, we will adopt the last method by assuming the classes follow a Gaussian distribution. This is motivated by the need to derive an analytically tractable and easy-to-implement reconstruction method. Under this assumption, $X \sim \mathrm{N}(\mu_{\mathrm{S}}, \Sigma_{\mathrm{S}})$ and $Y \sim \mathrm{N}(\mu_{\bar{\mathrm{S}}}, \Sigma_{\bar{\mathrm{S}}})$ where μ and Σ stand for the mean and covariance matrix, respectively. Since the two distributions are assumed linearly correlated, *x* and *y* may be linked by:

$$\boldsymbol{y} = \boldsymbol{K}\boldsymbol{x} + \boldsymbol{c} \tag{6.4}$$

$$\text{and} \quad \begin{cases} \boldsymbol{\mu_{\bar{S}}} = \boldsymbol{K}\boldsymbol{\mu_{S}} + \boldsymbol{c} \\ \boldsymbol{\Sigma_{\bar{S}}} = \boldsymbol{K}\,\boldsymbol{\Sigma_{S}}\,\boldsymbol{K^{T}} \end{cases} \tag{6.5}$$

where K is a transformation matrix, K^{T} its transpose, and c a bias vector. To estimate K and c, the Cholesky factorization is applied:

$$\begin{cases} \boldsymbol{c} = \boldsymbol{\mu}_{\bar{\mathrm{S}}} - K\boldsymbol{\mu}_{\mathrm{S}} \\ K = U_{\bar{\mathrm{S}}}\, V_{\mathrm{S}}^{-1} \end{cases}, \tag{6.6}$$

where $U_{\bar{\mathrm{S}}}$ and V_{S} are the lower and upper triangular Cholesky matrices related to the non-shadow and shadow classes, respectively. Once K and c are estimated, equation (6.4) is applied to compensate the pixels of the shadow class. Note that this process needs to be carried out for each couple of shadow and non-shadow classes.

When applying the shadow compensation, the restored area may appear noisy. This is due to the fact that the initial shadow distribution is much more concentrated than the non-shadow one. To mitigate this effect, the coefficient of variation (CV) defined as $\mathrm{CV}[i] = \sigma_i/\mu_i$ (σ_i and μ_i are the standard deviation and the mean along the i-th image band, respectively) is used to weigh the target variance of the reconstructed area [32]. In particular, we compute the CV ratio between the non-shadow and shadow classes, i.e., $\mathrm{N_{CV}}[i] = \mathrm{CV_{\bar{S}}}[i]/\mathrm{CV_{S}}[i]$, for the i-th original image band (i=1,...,N). Then, if $\mathrm{N_{CV}}[i] \geq 1$, the covariance matrix is corrected to reduce the variability of the non-shadow class:

$$cov_{NEW}[i,k] = \frac{cov[i,k]}{\mathrm{N_{CV}}[i] \cdot \mathrm{N_{CV}}[k]}. \quad \text{with} \quad (i,k) \in [1,N]^2 \tag{6.7}$$

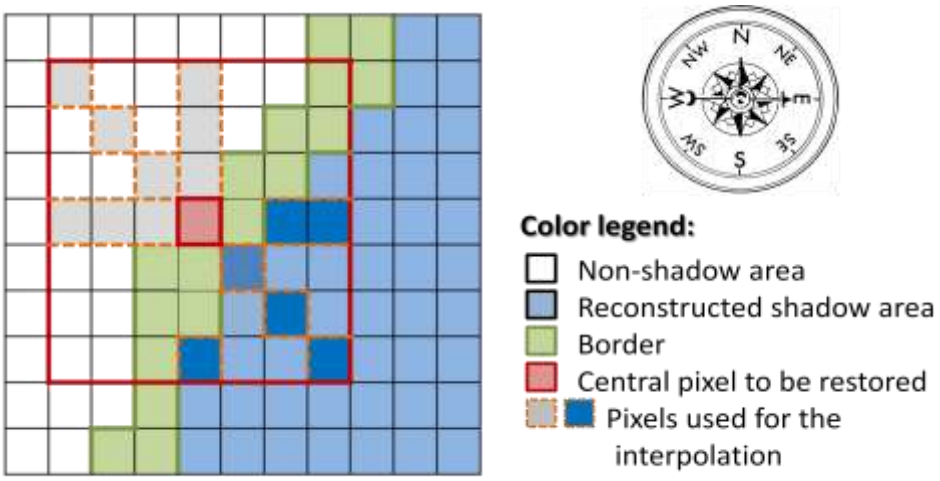


Figure 6.5 Illustration of the reconstruction of a border pixel with a 7 × 7 size window.

6.3.5. Border reconstruction

After the reconstruction of the shadow areas, the processing is not completely finished since thin borders between non-shadow and reconstructed shadow areas still remain with their original aspect, which may be in contrast with the two adjacent areas. In order to smooth such contrast, pixels of the borders undergo an easy-to-implement and fast contextual linear interpolation. In greater detail, a sliding window of predefined $S \times S$ size is adopted and within which four directional linear interpolations are considered. Among the following directions: North-South, West-East, NE-SW and NW-SE, just those crossing the given reconstructed shadow and non-shadow areas are retained. For instance, in Figure 6.5, only 3 directions are used to perform the linear interpolations (i.e., N-S, W-E and NW-SE). The NE-SW direction is removed since it does not cross the reconstructed shadow area within the window. Along a given direction, the linear interpolator is a simple affine transformation defined as follows:

$$z = m \cdot i + q, \tag{6.8}$$

where i stands for the coordinate along that direction while m and q are interpolation parameters. In order to estimate these lasts, we will make use of the least square estimator, also known under the name of pseudo-inverse method, which consists in our case in collecting the set of points which are: 1) within the window; 2) outside the border; and 3) along the considered direction crossing the central pixel of the window:

$$\begin{bmatrix} z_1 \\ \dots \\ z_N \end{bmatrix} = \begin{bmatrix} i_1 & 1 \\ \dots & 1 \\ i_N & 1 \end{bmatrix} \begin{bmatrix} m \\ q \end{bmatrix} \quad \Leftrightarrow Z = \Gamma \cdot \beta, \tag{6.9}$$

and then in getting the parameter estimate as follows:

$$\hat{\beta} = (\Gamma^T \Gamma)^{-1} \Gamma^T Z. \tag{6.10}$$

An interpolator is derived for each available direction. Different strategies could be adopted for merging the estimates provided by the obtained pool of interpolators. We found by visual inspection that the best fusion rule among four common rules, which are the MIN, MAX, average

and median rules, is achieved by retaining as estimate for the central pixel the largest value yielded by the pool (i.e., the MAX rule). This can be explained by the fact that the MAX rule provides values which are brighter and thus less contaminated by shadow or penumbra that may characterize undesired pixels not correctly detected during the border creation step and used in the interpolation process.

6.4. Experimental Results

6.4.1. Border reconstruction

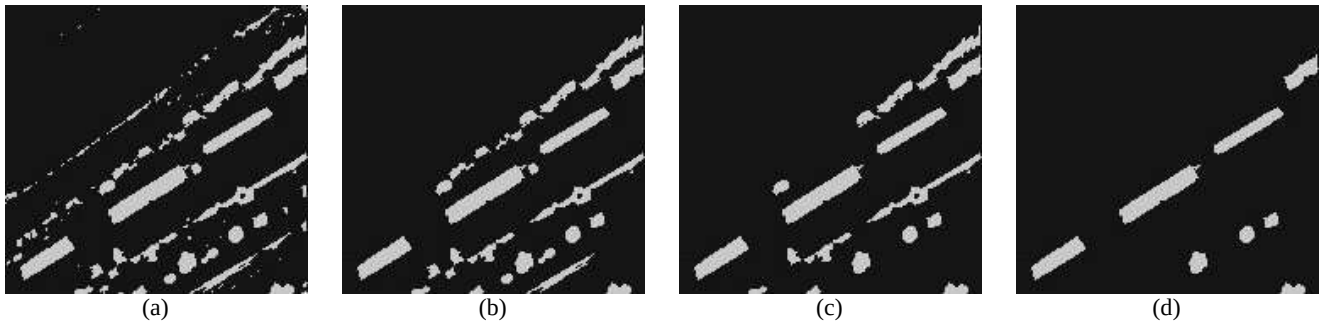
To evaluate the performance of the investigated method, three different images were used. They differ from the properties (distribution, size) of the shadows which depend on the kinds of shaded land covers and the acquisition time. The first is a QuickBird image of 450×600 pixel size (4 spectral bands, with a resolution of 0.6 meter) acquired on the 28th of February, 2008, and representing a part of the coastal region of Boumerdès (Algeria). It contains large shadows in a suburban area (see Figure 6.7(a)). The second and third images were acquired by IKONOS-2 (3 spectral bands, with a resolution of 1 meter). One represents the center of the city of Atlanta (USA), with a dimension of 420×500 pixels. It was taken in 1998 and contains long shadows in an urban area (see Figure 6.8(a)). The last image represents a peninsula of Jeddah (Saudi Arabia) with a dimension of 450×600 pixels and was acquired on the 11th of April 2004. It is dominated by the presence of green areas and exhibits only small shadow regions (see Figure 6.9(a)). From these images, different land covers, such as grass areas, roofs, parking lots, streets, sand and rock areas, as well as their corresponding shadows, were considered for generating the classification maps and thus the reconstructed images. The three images contain different amounts of shadow pixels. The percentages of shadow cover are: 16%, 42% and 15%, respectively.

6.4.2. Experimental setup and results

Before starting the reconstruction process, for each dataset, we prepared a ground-truth by photo-interpretation so that to generate the related mask and classification map conveying the shadow and non-shadow classes. The dominant non-shadow classes we defined for the Boumerdès image are: 1) vegetation ($\bar{S}_1$), 2) asphalt ($\bar{S}_2$), 3) bare soil ($\bar{S}_3$), and 4) sidewalk ($\bar{S}_4$). In the Atlanta image, the classes are: 1) bright roofs ($\bar{S}_1$), 2) asphalt ($\bar{S}_2$), 3) vegetation ($\bar{S}_3$), and 4) blue roofs ($\bar{S}_4$), while in the Jeddah image they correspond to: 1) vegetation ($\bar{S}_1$), 2) lane ($\bar{S}_2$), 3) asphalt ($\bar{S}_3$), and 4) roofs ($\bar{S}_4$). Afterwards, for completing the ground-truth, we defined the corresponding shadow counterparts (S_1, S_2, S_3, S_4) for each image. The numbers of training (TR) and test (TS) pixels adopted for the classification are given in Table 6.I. Note that: 1) for Jeddah a fifth non-shadow class, namely water, was defined but without its shadow counterpart; 2) the shadow counterparts of blue roofs and roofs are not available in the Atlanta and Jeddah images, respectively.

TABLE 6.I
NUMBER OF TRAINING (TR) AND TEST (TS) PIXELS USED TO CLASSIFY EACH CLASS OF THE THREE CONSIDERED IMAGES.

		Non-shadow classes				**Shadow classes**			
		$\bar{S}_1$	$\bar{S}_2$	$\bar{S}_3$	$\bar{S}_4$	S_1	S_2	S_3	S_4
Boumerdès	TR	893	263	912	478	174	1091	1493	597
	TS	372	184	334	216	161	477	764	194
Atlanta	TR	676	772	212	100	1740	1113	47	-
	TS	411	435	113	79	465	653	36	-
Jeddah	TR	1620	644	935	1592	276	8	63	-
	TS	678	330	226	709	67	3	25	-



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

Figure 6.6. Example of (a) initial mask image M_1, and post-processed mask by opening and closing M_2 obtained with different SE sizes: (b) 3×3, (c) 5×5, and (d) 7×7.

As described in the previous section, the first task of the proposed methodology is to generate a mask image useful to localize the shadow, the non-shadow and the border areas. For this purpose, a single level of wavelet decomposition was applied to extract wavelet features and, then, a binary SVM based on the Gaussian kernel was trained to recognize the group of shadow classes from the group of non-shadow classes. Note that the regularization and the kernel parameters of the SVM were tuned by cross-validation on the available training samples. Next, the SVM was applied on the whole image (composed from the original bands and the wavelet features) to yield a first mask M_1 (see Figures 6.7(b), 6.8(b) and 6.9(b)). A visual inspection of the masks suggests that, while for the Boumerdès and the Jeddah images the masks appear of good accuracy, the one for Atlanta suffers from a strong noise (very small shadow areas). Such noise is not due to the binary classifier but mainly to the presence of numerous cars, each characterized by its proper shadow. Since we think that very small shadow areas require a specific reconstruction process which is out of the scope of the present work, those are considered as noise. To reduce their presence in the mask, a morphological filtering is applied to M_1. Regarding the SE shape, we chose the popular square window [33]. In order to select the most appropriate SE size, we ran different experiments by adopting three values, namely 3×3, 5×5 and 7×7. As shown in Figures 6.6(b), 6.6(c) and 6.6(d), the smallest size (i.e, 3×3, see Figure 6.4(a)) appears visually the best one for preserving as much as possible the image details. Another motivation for using a 3×3 filter size is the absence of prior knowledge about the penumbra width. Indeed, such a filter size allows avoiding the removal of small width penumbra and handles reasonably large width penumbra. For these reasons, it has been adopted in all the following experiments. The outcome of this operation for the three images is provided in Figures 6.7(c), 6.8(c) and 6.9(c), respectively, which show a clear improvement of the

Table 6.II
LEGEND OF COLORS USED TO GENERATE THE CLASSIFICATION MAP FOR THE THREE CONSIDERED IMAGES.

Non-shadow classes				**Shadow classes**			
$\bar{S}_1$	$\bar{S}_2$	$\bar{S}_3$	$\bar{S}_4$	S_1	S_2	S_3	S_4

Table 6.III
CLASS ACCURACIES ACHIEVED ON THE TEST SAMPLES FOR EACH CLASS OF THE THREE CONSIDERED IMAGES.

	Non-shadow classes				**Shadow classes**			
	$\bar{S}_1$	$\bar{S}_2$	$\bar{S}_3$	$\bar{S}_4$	S_1	S_2	S_3	S_4
Boumerdès	100	100	100	100	100	92	96	60
Atlanta	99	98	91	99	93	92	0	-
Jeddah	100	100	85	91	96	0	52	-

Table 6.IV
USER'S AND PRODUCER'S ACCURACIES ACHIEVED FOR THE (A) BOUMERDÈS, (B) ATLANTA, AND (C) JEDDAH IMAGES.

	User's accuracy				**Producer's accuracy**			
	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4
Boumerdès	100	92	96	60	100	85	98	71
Atlanta	93	92	0	-	88	94	0	-
Jeddah	96	0	52	-	84	-	54	-

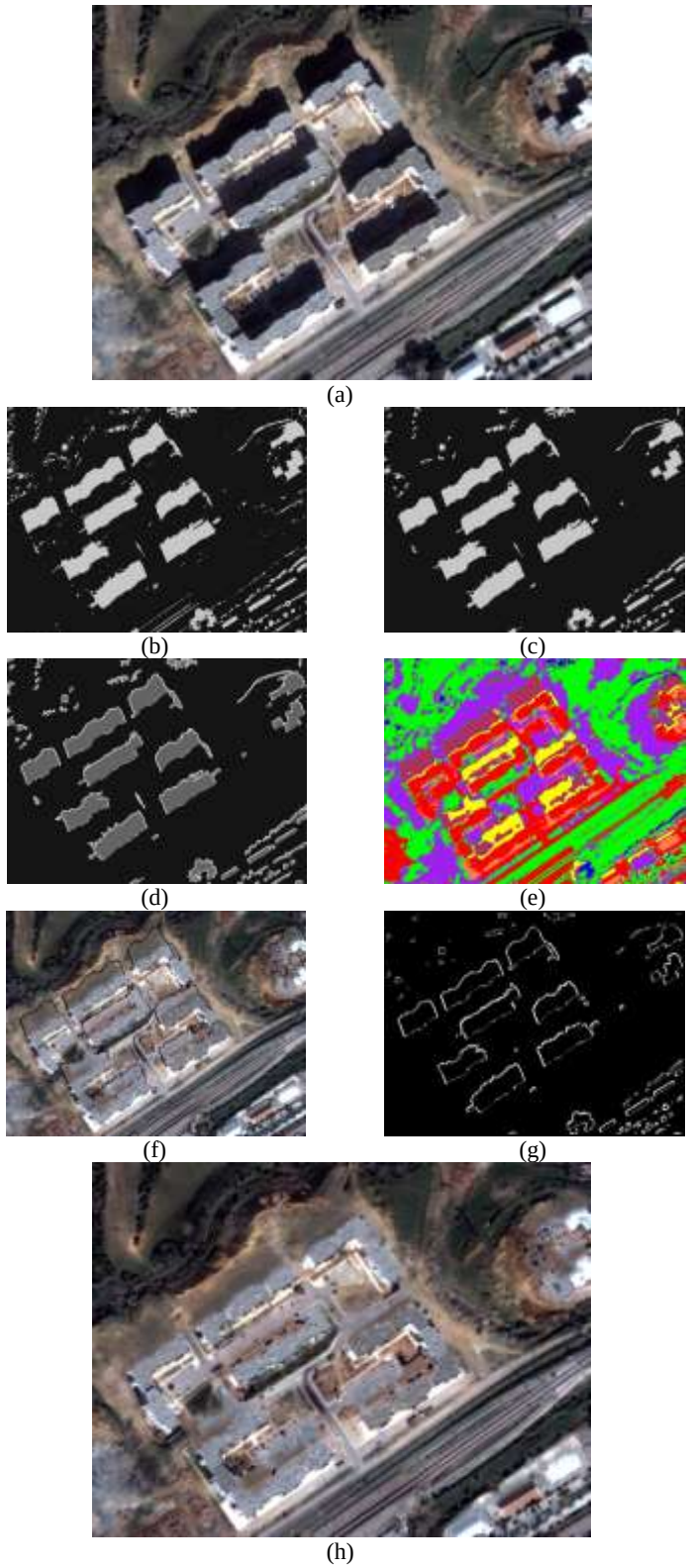
quality of the masks. In order to get from the masks M_2 the final masks M_{F}, the borders between the shadow and non-shadow classes were also produced to localize penumbras. For each image, a morphological filter was adapted according to the sun direction (see the SE in Figure 6.4(b) adopted for the Boumerdès image in Figure 6.3(c)), and then dilation and erosion operations were applied with it on M_2. The final result of the mask building phase is shown in Figures 6.7(d), 6.8(d) and 6.9(d). Such outcome is important since it allows to guide the successive multiclass classification and border reconstruction tasks.

Concerning the multiclass classifications, two multiclass SVMs with Gaussian kernel are trained on the basis of the available ground-truth, one for discriminating between the shadow classes and the other between the non-shadow classes. The resulting classification maps are provided in Figures 6.7(e), 6.8(e) and 6.9(e). The related legend of colors is provided in Table 6.II. The classification accuracies achieved on the test samples are listed in Table 6.III. It is interesting to observe that the accuracies are very high for the non-shadow classes. On an average over the three images, it is about 97%. As expected, for the shadow classes, results are less satisfactory since some classes exhibit poor accuracies (60% for the shaded sidewalks of the Boumerdès image and 52% for the shaded asphalts of the Jeddah image) while some have been completely misrecognized (case of shaded vegetation and shaded lanes for the Atlanta and Jeddah images, respectively). This puts under light the strong difficulty to correctly discriminate classes which by definition convey little proper information.

Before moving to the reconstruction step, as described in the previous methodological section, a quality check is performed in order to decide which shadow classes deserve to be reconstructed and which do not. In Table 6.IV, we report the user's and producer's accuracies for each shadow

class of each image. The quality threshold value we fixed is equal to 80%. This means that only if both user's and producer's accuracies are greater or equal to this value, the given shadow class will be reconstructed. Therefore, for the Boumerdès image, the shaded sidewalks were not a candidate for reconstruction. For the Atlanta image, two shadow classes among four underwent the same fate, while for the Jeddah image it was possible to reconstruct just the shaded vegetation areas.

Finally, for each image and each type of shadow area which successfully passed through the previous quality check, a linear compensation is applied (see Figures 6.7(f), 6.8(f) and 6.9(f)) and the corresponding borders are interpolated (see Figures 6.7(g), 6.8(g) and 6.9(g)), as explained in subsections 6.3.4 and 6.3.5, respectively. The final reconstruction is provided for the three images in Figures 6.7(h), 6.8(h) and 6.9(h), respectively. In greater detail, focusing first on the Boumerdès image (see Figure 6.7(h)), one can notice that some areas, mostly small, still remain dark (shaded) since they have not been reconstructed. Indeed, these areas correspond to the shaded sidewalks which represent a shadow class rejected by the quality check. We recall that this last is based on the principle that "better leave things as they are than commit errors". By visual inspection, the remaining reconstructed shadow areas look very realistic at the point that in some cases it is difficult to discern them with respect to the non-shadow areas. This means that for this image the reconstruction process was capable to reproduce satisfactorily not only the spectral properties of the shadow areas but also their textural ones. In the Atlanta image (see Figure 6.8(h)), results are more mitigated, also because of the darkness and the heterogeneity of the shadow regions, especially for the asphalt class in the parking. This is explained in part by the fact that some areas of the image are misclassified and thus reconstructed by error. Using a target detection terminology, such issue could be defined as a problem of false alarms (i.e., non-shadow areas classified as shadows). An example is a dark roof of a building in the top of the image which is classified as a shadow class and, therefore, reconstructed by error. A second example is the self shadows of some buildings on the left part of the image which are sometimes recognized as shadows and sometimes not. A third example is the shadow on the roof (in the center of the image) reconstructed as asphalt. Another problem is related to the intrinsic complexity of some classes. Indeed, if we consider the class asphalt, which is characterized by a high classification accuracy as well as its shadow counterpart, we observe that its reconstructed shadows appear noisy. The cause is the multimodal nature of this class, which actually spans spectrally heterogeneous objects such as streets and parking lots. Note however that some shadow regions are well reconstructed, like the bright roofs in the bottom part of the image. The Jeddah image is instead marked by a lot of thin shadows mostly located in vegetation areas. The shadow thinness is explained by the steepness of the sun angle at the local image acquisition hour (11:17 a.m.). For this image, the reconstruction, which was limited to the shaded vegetation, was globally satisfactory (see Figure 6.9(h)).



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 6.7. Reconstruction results for Boumerdès image. (a) Original image. (b) Binary mask. (c) Post-processed mask. (d) Mask with borders. (e) Multiclass classification. (f) Shadow reconstruction. (g) Border interpolation. (h) Final output image.

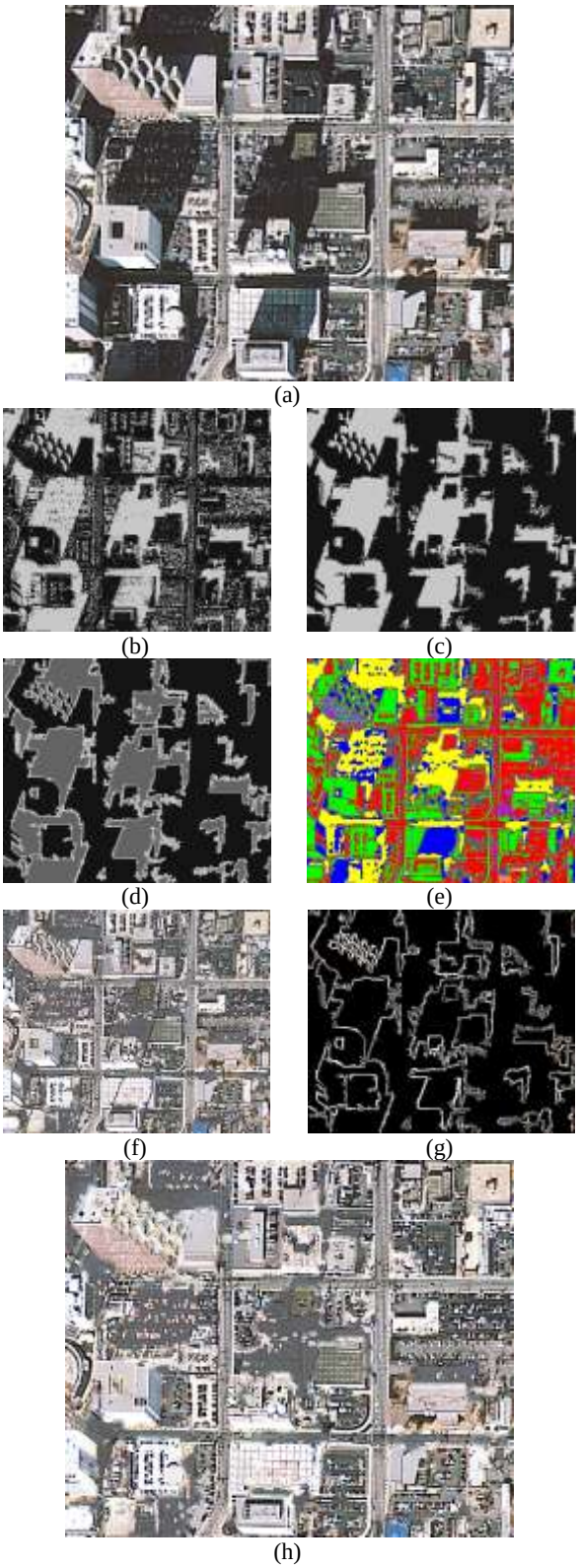


Figure 6.8. Reconstruction results for Atlanta image. (a) Original image. (b) Binary mask. (c) Post-processed mask. (d) Mask with borders. (e) Multiclass classification. (f) Shadow reconstruction. (g) Border interpolation. (h) Final output image.

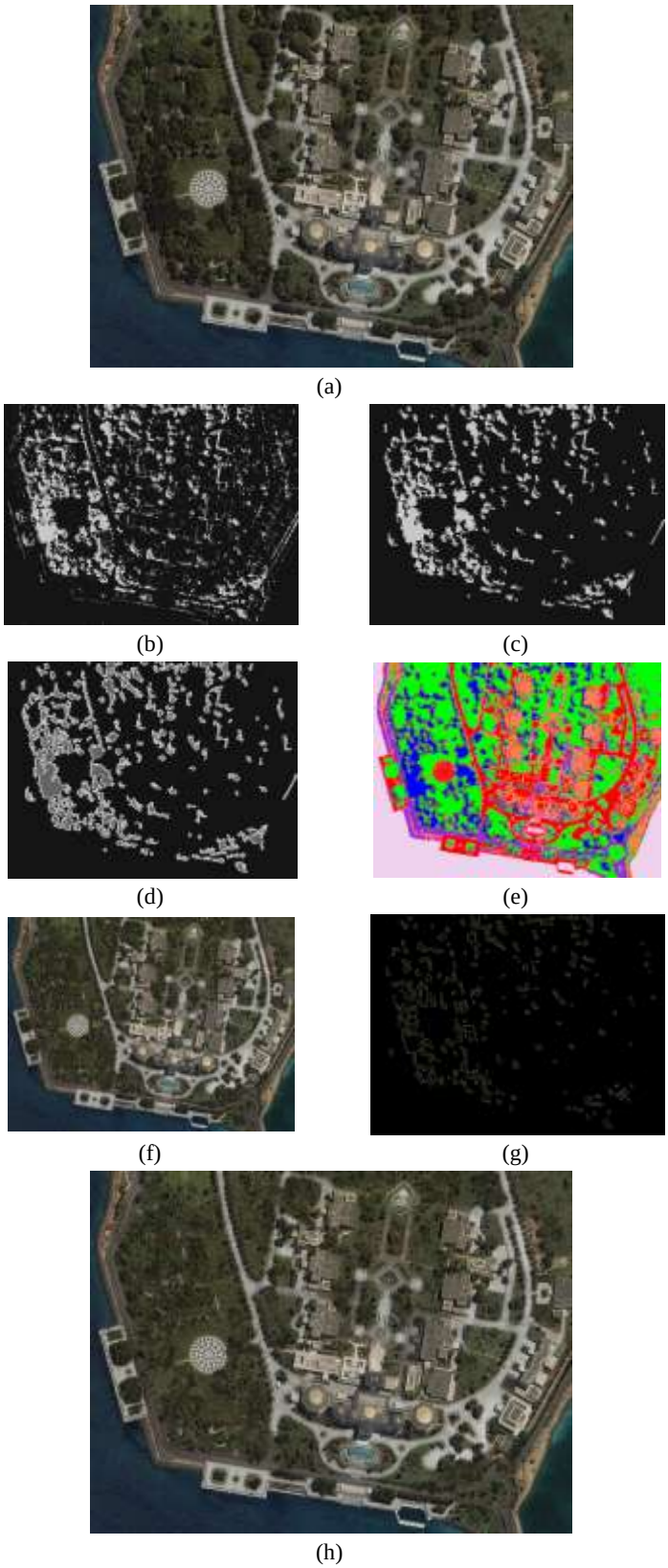


Figure 6.9. Reconstruction results for Jeddah image. (a) Original image. (b) Binary mask. (c) Post-processed mask. (d) Mask with borders. (e) Multiclass classification. (f) Shadow reconstruction. (g) Border interpolation. (h) Final output image.

6.4.3. Reconstruction impact on classification accuracy

Since the production of classification maps represents one of the most widespread applications of remote sensing images, we think it also important to evaluate the quality of the reconstruction process in terms of classification accuracy. In particular, we evaluate the classification accuracy before and after the application of our reconstruction methodology. In the ideal case, it is expected that all areas of shadow classes are classified correctly as the corresponding non-shadow classes.

After the compensation of the shadow classes in the images, we reclassified the resulting reconstructed images, with the same classifier adopted above. From the maps shown in Figure 6.10, we observe that most of the compensated shadow classes have been classified as the corresponding non-shadow classes. However, some shadow areas still remain. Two explanations can be found: 1) either the shadow areas have not been correctly compensated; or 2) they have been misclassified (in the second classification round). Such a visual analysis of the classification maps is confirmed from a quantitative viewpoint. Indeed, Tables 6.V and 6.VI show the benefit of applying the reconstruction process on the classification maps, namely: 1) much less areas of the shadow classes survive and a significant part of them is assigned to the correct non-shadow class (Table 6.V); 2) this results in a more complete (since less shadow areas are present) and more accurate classification map (Table 6.VI).

TABLE 6.V
PERCENTAGE OF SHADOW PIXELS CORRECTLY CLASSIFIED AFTER SHADOW RECONSTRUCTION.

	$S_1 \rightarrow \overline{S}_1$	$S_2 \rightarrow \overline{S}_2$	$S_3 \rightarrow \overline{S}_3$	$S_4 \rightarrow \overline{S}_4$
Boumerdès	76	95	74	0
Atlanta	81	78	15	-
Jeddah	58	0	55	-

TABLE 6.VI
CLASSIFICATION USER ACCURACIES (IN PERCENT) BEFORE AND AFTER SHADOW RECONSTRUCTION. THEY ARE COMPUTED BY MERGING NON-SHADOW AND SHADOW AREAS INTO THE SAME THEMATIC CLASS. THIS LAST IS CORRECTLY RECOGNIZED IF ALL THE PIXELS ARE CLASSIFIED AS NON-SHADOW.

	Before reconstruction				**After reconstruction**			
	$\overline{S}_1 \cup S_1$	$\overline{S}_2 \cup S_2$	$\overline{S}_3 \cup S_3$	$\overline{S}_4 \cup S_4$	$\overline{S}_1 \cup S_1$	$\overline{S}_2 \cup S_2$	$\overline{S}_3 \cup S_3$	$\overline{S}_4 \cup S_4$
Boumerdès	97	87	85	73	100	100	97	99
Atlanta	92	91	50	100	100	71	75	98
Jeddah	93	95	64	88	99	100	74	88

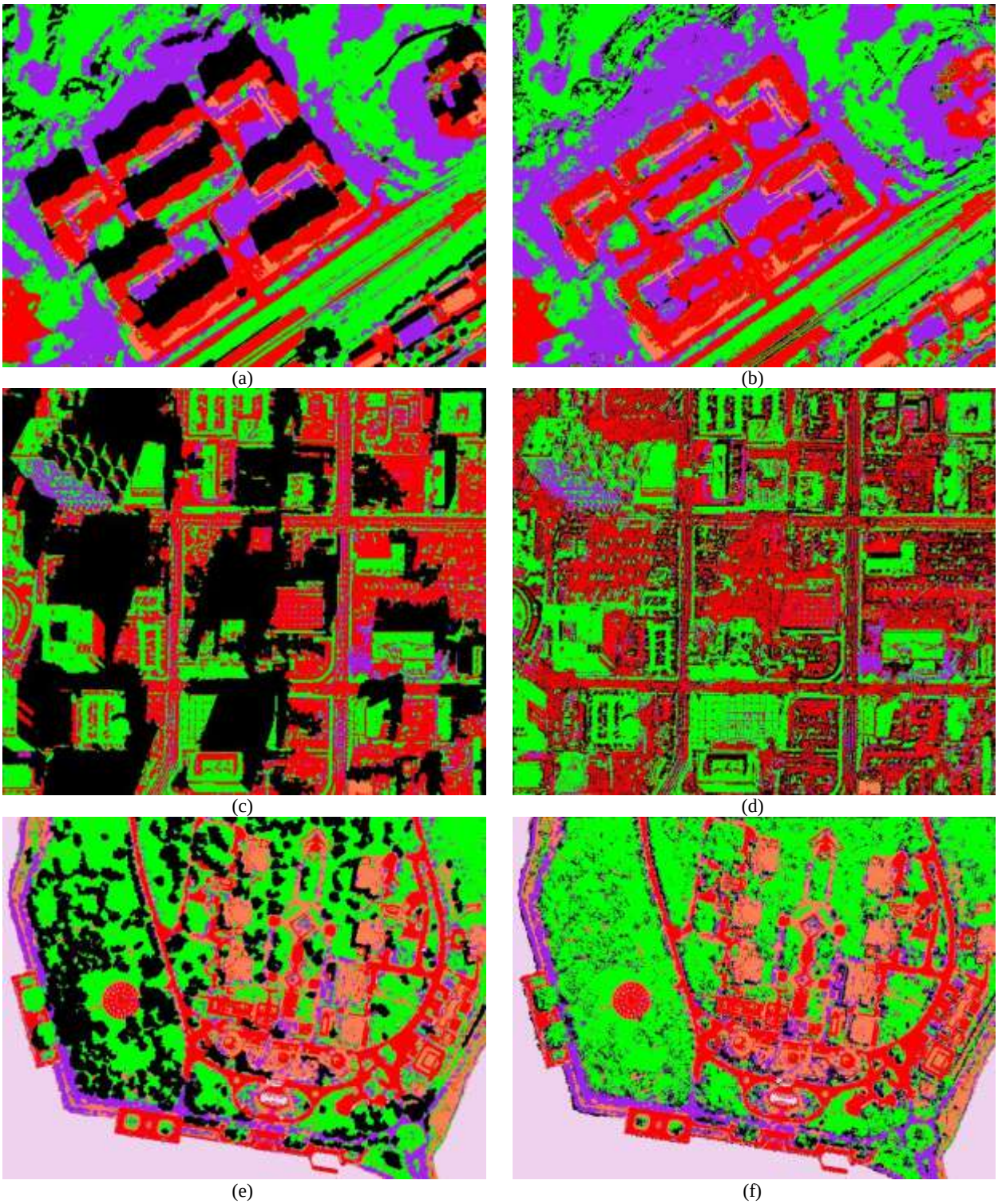


Figure 6.10. Classification results before and after shadow reconstruction: (a) and (b) Boumerdes image, (c) and (d) Atlanta image, and (e) and (f) Jeddah image. All shadow classes are grouped in a unique class (in black).

6.4.4. Comparative analysis

In this subsection, we compare our methodology with two other recently published techniques proposed to remove shadow effects from remote sensing imagery. The first one implements a

normalized saturation-value difference index (NSVDI) in hue-saturation-value (HSV) color space for the detection of shadow regions [12]. To restore these obscured regions, it applies a histogram matching method. The second technique uses a panchromatic (PAN) image to manually threshold shadow pixels by visually inspection [20]. In order to recover shadow values, it uses a linear regression in each multispectral band. Both techniques select the best threshold by visual inspection. Note that in these techniques, the authors do not implement any specific strategy for dealing with the borders between shadow and non-shadow regions. Because of this limitation, in our comparison, we will ignore the differences at the borders. Figures 6.11(a), 6.11(c), and 6.11(e) show the binary masks obtained with our methodology and the two other techniques on the Atlanta image (i.e., the most complex image considered in our experiments), while Figures 6.11(b), 6.11(d), and 6.11(f) depict the corresponding reconstructed images. Note that the binary masks in Figures 6.11(c) and 6.11(e) are characterized by a "salt and pepper" effect. This is explained by the fact that binarization is yielded by simple thresholding, while in our method it is obtained by a more sophisticated classification procedure. As a direct consequence, the corresponding compensation results are of low quality.

To improve the results of these two methodologies, we integrated our morphological approach so as to clean the binary masks. The obtained enhanced images are shown in Figure 6.12. Comparing the images generated with [12] and [20], it seems that the first one reconstructs a bit better the asphalt class, but both recover a brighter and noisier image compared with our result. In addition, the bright roofs class seems better reconstructed by our methodology with respect to the reference methods. Note that in all the three techniques we still have reconstruction errors with the dark roof of a building in the top of the image, considered as shadow. Self-shadows of some buildings in the left part of the image are in some cases still recognized as shadow. The roof in the center of the image is reconstructed as asphalt in all the cases, and with [20] the remaining part of the roof is also reconstructed because misclassified as shadow.

6.5. Conclusion

This paper deals with the important and challenging problem of reconstruction of VHR images obscured by the presence of shadows. The proposed methodology is supervised. The shadow areas are not only detected but also classified so that to allow their customized compensation. The classification tasks are implemented by means of the state-of-the-art SVM approach. A quality check mechanism is integrated in order to limit misreconstruction problems. Moreover, borders are explicitly handled by adaptive morphological filters and linear interpolation for the prevention of possible border artifacts in the reconstructed image.

In general, from the obtained results, different considerations may be deduced:

1) The proposed methodology can yield visually realistic shadow-free images with a promising preservation of the spectral and textural properties of the obscured objects. Moreover, it improves the results in terms of classification accuracy.
2) Because of the quality check mechanism, not all shadow areas are systematically reconstructed. Part of them may remain unchanged due to the difficulty to recognize them correctly or because of their small size.

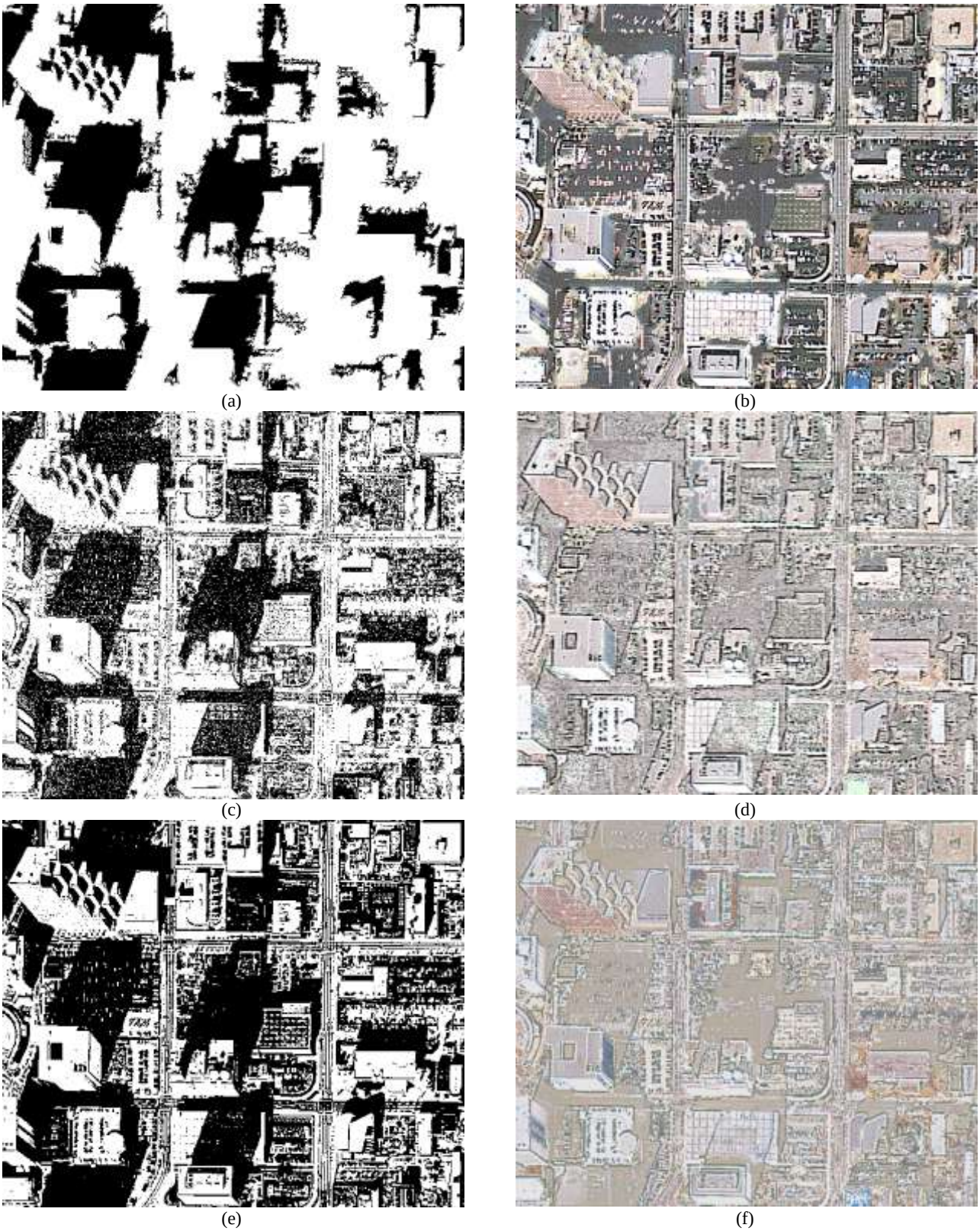
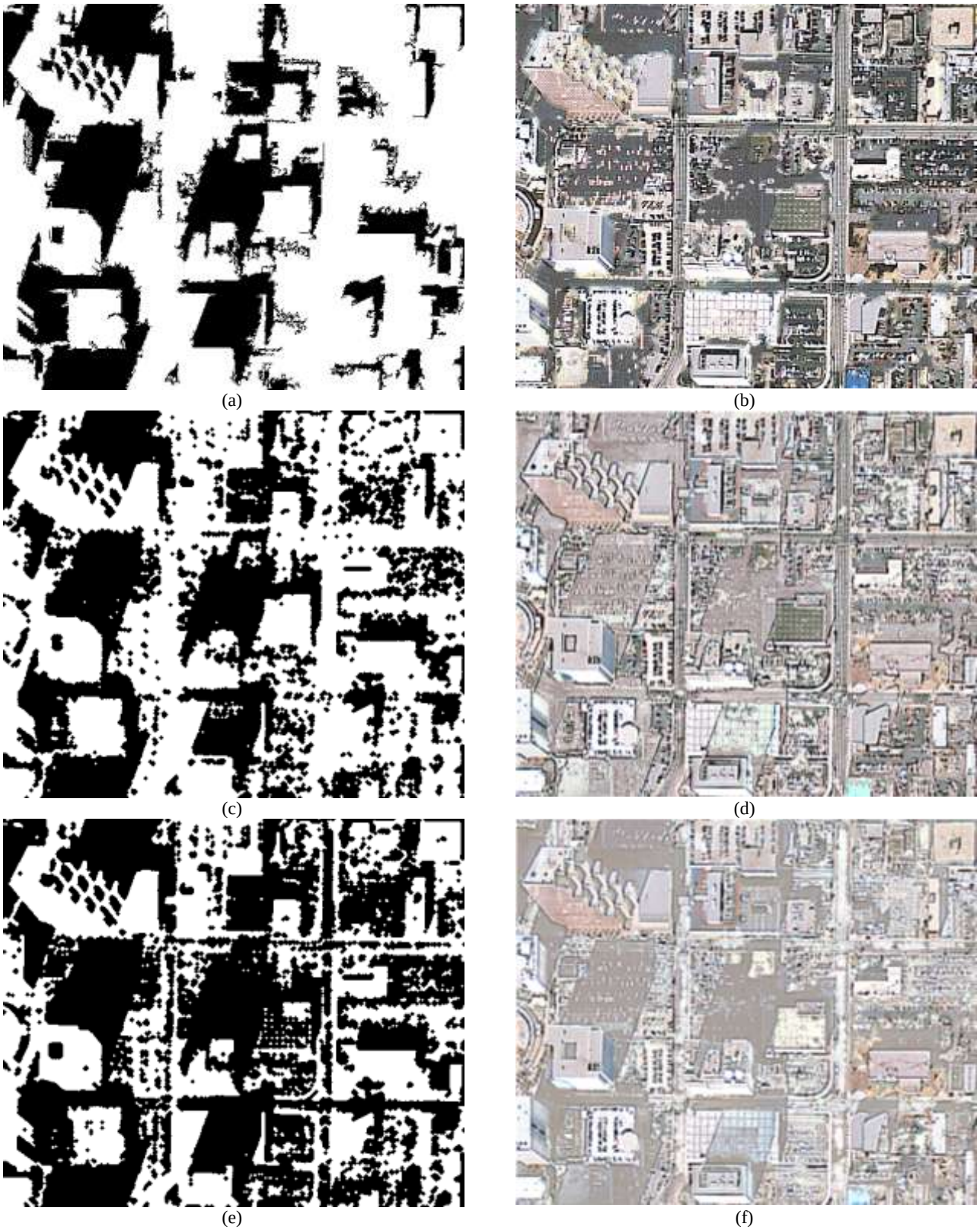


Figure 6.11. Comparison with methods [12] and [20] for Atlanta image. (a), (c) and (e) binary masks and (b), (d) and (f) reconstructed images associated with our method, [12] and [20], respectively.

3) The recognition of the couples of non-shadow and shadow classes can help in a more accurate compensation but involves a classification process which can create false shadow areas and thus overload the image with misreconstructions.



(a) (b) (c) (d) (e) (f)

Figure 6.12. Comparison with methods [12] and [20] enhanced with the morphological approach for Atlanta image. (a), (c) and (e) binary masks and (b), (d) and (f) reconstructed images associated with our method, [12]and [20], respectively.

4) In case where non-shadow classes are multimodal, the implemented linear compensation method can be found inappropriate. Nonlinear alternatives could be required but at the cost of a higher computational burden.

5) The comparison with two reference techniques for shadow compensation points out that our methodology returns better reconstructions.

This work has the merit of facing the problem as a whole. The proposed processing chain relies on different advanced image processing and pattern recognition technologies. Though interesting results have been obtained, still further methodological improvements are required including: i) reconstructability evaluation in order to a priori assess if a particular shadow area can potentially be reconstructed or not; ii) active learning process to adapt as most as possible the ground-truth to the image reconstruction requirements; iii) reinforce the quality check with a multimodality assessment either to reject complex classes or to operate with a multimodal nonlinear compensation; iv) accompanying the reconstruction map with a kind of confidence map as done for the problem of cloud-contaminated images [34].

Additional future directions could be envisioned. First, the problem of the SE shape and size could be faced by means of an automatic adaptation procedure according to the sensor resolution and the penumbra width, which depends on the sun direction and the building heights. Second, since the reconstruction of shadow regions strongly depends on the accuracy of the classification maps, the height derived from a digital elevation model could be considered as an additional input feature to better discriminate between the thematic classes. Finally, a third future direction could be to face the reconstruction problem with more sophisticated statistical models. Though they would increase the computational complexity, they would lead to a better fitting of the shadow and non-shadow classes, thus resulting in a potentially better reconstruction quality.

6.6. Acknowledgment

The authors would like to thank C. C. Chang and C. J. Lin for supplying the software library LIBSVM used in this research (http://www.csie.ntu.edu.tw/~cjlin/libsvm). The authors are grateful to CNES for making available the IKONOS-2 Atlanta image through the Orfeo initiative and to Dr. Y. Bazi (King Saud University, Saudi Arabia) for providing the Jeddah image.

6.7. References

[1] T. Kim, T. Javzandulam and T-Y. Lee, "Semiautomatic reconstruction of building height and footprints from single satellite images," in *Proc. IGARSS*, Barcellona, Spain, Jul. 2007, vol. 2, pp. 4737-4741.

[2] A. Massalabi, H. Dong-Chen, G.B. Benie and E. Beaudry, "Detecting information under and from shadow in panchromatic IKONOS images of the city of Sherbrooke," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 3, pp. 2000–2004.

[3] P.M. Dare, "Shadow analysis in high-resolution satellite imagery of urban areas," *J. of the American Society for Photogrammetry and Remote Sensing*, vol. 71, pp. 169-177, 2005.

[4] K. Kouchi and F. Yamazaki, "Characteristic of tsunami-affected areas in moderate-resolution satellite images," *IEEE Trans. Geosci. and Remote Sens.*, vol. 45, no. 6, pp. 1650-1657, May 2007.

[5] J. Liu, J. Yang and T. Fang, "Color property analysis of remote sensing imagery," *J. Act Photonica Sinica*, vol. 38, no. 2, pp. 441-446, Feb. 2009.

[6] G. Finlayson and S. Süsstrunk, "Optimization for hue constant RGB sensors," in *Proc. IS&T/SID*, Scottsdale, AZ, USA, Nov. 2002, vol.10, p.343-348.

[7] E. Salvador, A. Cavallaro and T. Ebrahimi, "Shadow identification and classification using invariant color models," in *Proc. ICASSP*, Salt Lake City, UT, USA, May 2001, vol. 3, pp. 1545-1548.

[8] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 9, pp. 62-69, Jan. 1979.

[9] V. Tsai, "A comparative study on shadow compensation of color aerial images in invariant color models," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no.6, pp. 1661−1671, Jun. 2006.

[10] K. L. Chung, Y. R. Lin, and Y. H. Huang, "Efficient shadow detection of color aerial images based on successive thresholding scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 671−682, Feb. 2009.

[11] D. Cai M. Li, Z. Bao, Z. Chen, W. Wei, and H. Zhang, "Study on shadow detection method on high resolution remote sensing image based on HIS space transformation and NDVI index," in *Proc. International Conference on Geoinformatics*, Beijing, China, Jun. 2010, pp.1–4.

[12] H. Ma, Q. Qin, and X. Shen, "Shadow segmentation and compensation in high resolution satellite images," in *Proc. IGARSS*, Boston, MA, USA, Jul. 2008, vol. 2, pp. 1036–1039.

[13] H.Y. Yu, J.G. Sun, L.N. Liu, Y.H. Wang, and Y.D. Wang, "MSER based shadow detection in high resolution remote sensing image," in *Proc. ICMLC*, Qingdao, China, Jul. 2010, pp. 780–783.

[14] S. Wang, and Y. Wang, "Shadow detection and compensation in high resolution satellite images based on retinex," in *Proc. ICIG*, Xi'an, China, Sep. 2009, pp. 209–212.

[15] A. Makarau, R. Richter, R. Müller and P. Reinartz, "Adaptive shadow detection using a blackbody radiator model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2049-2059, Jun. 2011.

[16] P. Sarabandi, F. Yamazaki, M. Matsuoka, and A. Kiremidjian, "Shadow detection and radiometric restoration in satellite high resolution images," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 6, pp. 3744–3747.

[17] A. Massalabi, H. Dong-Chen, G.B. Benie and E. Beaudry, "Detecting information under and from shadow in panchromatic IKONOS images of the city of Sherbrooke," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 3, pp. 2000−2004.

[18] J. Su, X. Lin, and D. Liu, "An automatic shadow detection and compensation method for remote sensed color images," in *Proc. ICSP*, Beijing, China, Nov. 2006, vol. 2.

[19] T. Nakajima, G. Tao, and Y. Yasuoka, "Simulated recovery of information in shadow areas on IKONOS image by combining ALS data," in *Proc. ACRS*, Kathmandu, Nepal, 2002.

[20] F. Yamazaki, W. Liu, and M. Takasaki, "Characteristic of shadow and removal of its effects for remote sensing imagery," in *Proc. IGARSS*, Cape Town, South Africa, Jul. 2009, vol. 4, pp.426–429.

[21] T. Kasetkasem and P.K. Varshney, "An optimum land cover mapping algorithm in the presence of shadow," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, Jun. 2011.

[22] B. Abdel Latif, R. Lecerf, G. Mercier, and L. Hubert-Moy, "Preprocessing of low-resolution time series contaminated by clouds and shadows," *IEEE Trans. Geosci. and Remote Sens.*, vol. 46, no. 7, pp. 2083-2096, Jul. 2008.

[23] P. Soille, *Morphological Image Analysis*. Springer ed., 1999.

[24] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. and Remote Sens.*, vol. 41, no. 9, pp. 1940-1949, Sep. 2003.

[25] N. Ghoggali and F. Melgani, "Genetic SVM approach to semisupervised multitemporal classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 212-216, Apr. 2008.

[26] E. Pasolli, F. Melgani, and M. Donelli, "Automatic analysis of GPR images: A pattern recognition approach," *IEEE Trans. Geosci. and Remote Sens*., vol. 47, no. 7, pp. 2206-2217, Jul. 2009.

[27] T. Habib, J. Inglada, G. Mercier, J. Chanussot, "Support vector reduction in SVM algorithm for abrupt change detection in remote sensing," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 606-610, Jul. 2009.

[28] R. Singh, R.E. Vasquez and R. Singh, "Comparison of Daubechies, Coiflet, and Symlet for edge detection," in *Proc. SPIE Visual Information Processing VI*, Orlando, FL, USA, Apr. 1997, vol. 3074, pp. 151–159.

[29] P. Soille and M. Pesaresi, "Advances in mathematical morphology applied to geosciences and remote sensing," *IEEE Trans. Geosci. and Remote Sens*., vol. 40, pp. 2042-2055, Sep. 2002.

[30] A. Singh, U. Ghanekar, C. Kumar, and G. Kumar, "An efficient morphological salt-and-pepper noise detector," *Int. J. Advanced Networking and Applications*, vol. 2, no. 5, pp. 873-875, 2011.

[31] R. Kohavi and F. Provost, "Glossary of Terms," in Machine Learning, Kluwer Academic Publishers, Boston, vol. 30, no. 2-3, pp. 271-274, 1998.

[32] W. A. Hendricks and K. W. Robey, "The sampling distribution of the coefficient of variation," *Annals of Mathematical Statistics*, vol. 7, pp. 129-132, 1936.

[33] R. M. Haralick, S. R. Sternberg and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 523-550, Jul. 1987.

[34] S. Benabdelkader and F. Melgani, "Contextual spatiospectral postreconstruction of cloud-contaminated images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 204–208, Apr. 2008.

7. Assessing the Reconstructability of Shadow Areas in VHR Images

Abstract *– Very high resolution (VHR) images are appreciated for their high-level details which significantly increase their application potential. However, typically, VHR images are affected by the presence of shadows. An attempt solution to the problem of shadows is to restore shadow-contaminated regions, by compensating the value of shaded pixels. Unfortunately, it may happen that not all shadow areas are possible to restore. In this chapter, we propose different criteria useful to help in understanding a priori if it is possible or not to reconstruct a specific shadow area. An ideal reconstructability criterion should not tolerate that an unreconstructable shadow area is assigned as reconstructable and, at the same time, should maximize the probability of detection of reconstructable areas. Several evaluation criteria working at the pixel and textural levels are presented. Furthermore, in order to select the best criteria, a fuzzy logic combination of the criteria is explored. A thorough experimental analysis is reported and discussed. It leads to the definition of a final global index based on the fusion of two single criteria which are the Kullback-Leibler divergence and the angular second-moment difference.*

The work presented in this chapter has been accepted to be published in the *IEEE Trans. Geosci. Remote Sens.*; Co-authors: F. Melgani, G. Mercier, Y. Bazi.

7.1. Introduction

The presence of shadows in very high resolution (VHR) optical images represents an important and inevitable obstacle for their full exploitation. Although it is feasible to exploit shadow characteristics to estimate the height of buildings for instance, usually shadows are viewed as an undesired presence incurring in a partial or total loss of information. Mainly for this reason shadow-contaminated images need preferably to be compensated (reconstructed). In general, after the detection of shadow regions, reconstruction techniques try to remove these unwanted presences by restoring the obscured pixels in order to obtain an enhanced image, free of shadows. Unfortunately, it is not always possible to compensate shadow regions in a satisfactorily way. Shadow reconstruction is different from the cloud reconstruction problem since the former is typically viewed as a partial contamination problem while the latter is reduced to a completely missing data issue [1]-[4].

In the literature, there exist essentially three methods to recover shadow regions: 1) gamma correction; 2) histogram matching; and 3) linear correlation [5]. If it is supposed that the surface texture does not change dramatically when it is shadowed, shadow removal can be performed by the nonlinear gamma correction (*GC*) method. Indeed, in [6], the authors applied a contextual texture analysis between a segment of shadow and its neighbors, and, knowing the relationship between the two areas, they restore the intensity value of the shadow pixels. In [7], the shadow areas are detected automatically and in a second step, the initial RGB satellite image is projected to the hue, intensity and saturation (HIS) color space, where shaded pixels are more easily detected. The authors adopt a histogram matching (*HM*) method to restore shadow pixels. In [8], the height data of an airborne laser scanner (ALS) are exploited to recover the shadow position in an IKONOS image. In a second step, this information is used to first overlay and second eliminate the shadow presence. In this case, the authors compare the results of two solutions: gamma correction and linear transformation (*LT*). If the shadow restoration process mainly depends on the spectral signature of the image spectral bands, the authors of [9] find convenient to first threshold and then apply a linear regression separately in each spectral band. Recently, a complete processing chain based on a support vector machine (SVM) regression was developed for the detection and reconstruction of shadow in VHR images [10]. Another potentially direction, is to exploit physical properties (e.g., temperature) of a black-body radiator to discriminate the presence of shadow [11]. A different approach to restore shadow pixels start from the idea to compensate the atmospheric effects; namely to remove sun effects on the radiance image and work in a compensated reflectance image, obtained after an ICARE (Cloud-Aerosol-Water-Radiation Interactions) processing which separates irradiance and radiance components [12]. Adopting a digital elevation model (DEM), the authors obtain accurate results, except in the presence of vegetation classes, where the DEM results less accurate.

In this chapter, our objective is to try to answer to the following question: *Is it possible to know a priori if a shadow area can be well recovered?* The core of the present contribution is to propose several criteria, capable to estimate *a priori* if it is possible or not to restore a specific shadow class, independently from the adopted reconstruction method. They work at the pixel-level, or at the textural-level, namely: histogram quantization error, gray level ratio, two-sample

Kolmogorov-Smirnov test, variance ratio, negentropy, Kullback-Leibler divergence, angular second-moment difference and homogeneity distance. Furthermore, in order to select the best criteria, a fuzzy logic combination of the criteria is explored.

The remainder of this chapter is organized as follows. In Section 7.2, five common shadow reconstruction methods are reviewed. Section 7.3 formulates the reconstructability problem. Section 7.4 describes the proposed evaluation criteria and Section 7.5 presents the criterion selection and the decision making procedure. Section 7.6 reports experimental results and Section 7.7 draws the conclusions.

7.2. Shadow Reconstruction Methods

Five supervised shadow reconstruction methods are briefly described in this section. They will be utilized in the experimental part to assess the validity of the proposed criteria. Three of them are state-of-the-art techniques, while the other two are improved methods. In the following, it is supposed that, in a given image *I*, shadows are categorized according to the class (ground-cover type) they obscure. Moreover, it will be assumed that X is a random variable (RV) characterizing the shadow area, with sample x being scalar and Y the RV characterizing the non-shadow area (belonging to the same class as the shadow area) with sample y (a scalar value).

7.2.1. Linear transformation (LT)

The *LT* method can be applied to perform a supervised linear regression to restore shadow pixels. Starting from the shadow area, this method restores shaded pixels making them as more similar as possible to the corresponding non shadow pixels, by adjusting the shift and the scale of the two respective distributions [9]. This regression is based on the mean μ and on the variance σ^2 values of the classes. Indeed, assuming the RV of shadow class as $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$ and the corresponding RV of non-shadow class as $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$, the reconstruction of the shadow class can be reduced to a simple RV transformation:

$$X \sim \mathrm{N}(\mu_X, \sigma_X^2) \rightarrow \hat{Y} \sim \mathrm{N}(\mu_Y, \sigma_Y^2). \tag{7.1}$$

This technique supposes that shadows are a linear noise that can be removed using the following simple relationship:

$$\hat{y} = \frac{\sigma_Y}{\sigma_X}(x - \mu_X) + \mu_Y \,. \tag{7.2}$$

7.2.2. Histogram matching (HM)

HM is a method generally used for image registration. It exploits cumulative distribution functions (CDF) of two different RVs (see Figure 7.1) [13]. Starting from a pixel value x, it is possible to reach the corresponding quantile value q of $\mathrm{CDF}_X(x)$. From q of $\mathrm{CDF}_X(x)$, the method reaches the quantile value q of $\mathrm{CDF}_Y(y)$, and then the corresponding y pixel value. This method

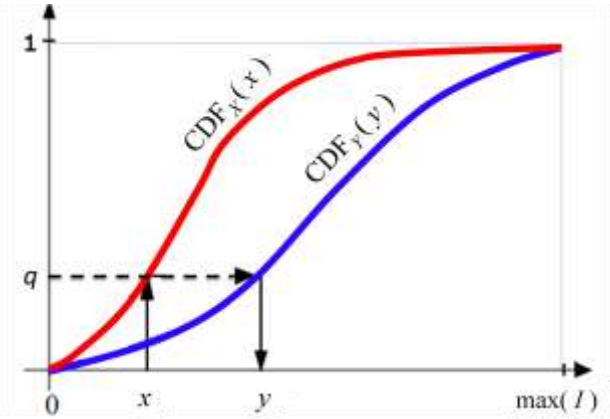


Figure 7.1. Illustration of the histogram matching method. **max**(***I***) corresponds to the max value in image *I*.

can be applied to restore the shadow class, as in [7], where the $\mathrm{CDF}_X(x)$ and $\mathrm{CDF}_Y(y)$ correspond respectively to the shadow and non-shadow CDFs:

$$\hat{y} = \mathrm{CDF}_Y^{-1}\big(\mathrm{CDF}_X(x)\big)\,. \tag{7.3}$$

Several techniques may be used for estimating the CDFs, we will focus on three of them.

7.2.2.1. Rank statistics (HM_{rank})

It is known that the estimation of the CDF of a RV X may be performed from the ranking of its samples [14]. From a set of samples $\{x_1, x_2, \dots, x_n\}$, we can define the ordered samples $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(q)} \leq \cdots \leq x_{(n)}$. Accordingly, $q = \mathrm{CDF}_X(\,x_{(q)}\,)$.

7.2.2.2. Mean of grayscale ratio (HM_{MGR})

To take into consideration the fact that the support of X (the shadow area) is (much) lower than the support of the non-shadow area Y, it may be useful to perform the estimation of those pdfs by using histograms having the same bin width. This yields a histogram comparison with one of them (the non-shadow one) being coarsely estimated as shown on Figure 7.2. Then Equation (7.3) is applied as it stands.

7.2.2.3. Roulette of the grayscale ratio (HM_{RGR})

The *RGR* is also a proposed improved method of the *HM*, which works in a similar way as the previous method (*MGR*). This previous technique may be thought of as a quantization of the quantile values of CDF_Y. With *RGR*, it may be thought of as a de-quantization by choosing randomly an original-resolution value q and then y, and not a coarse one. A roulette wheel, which

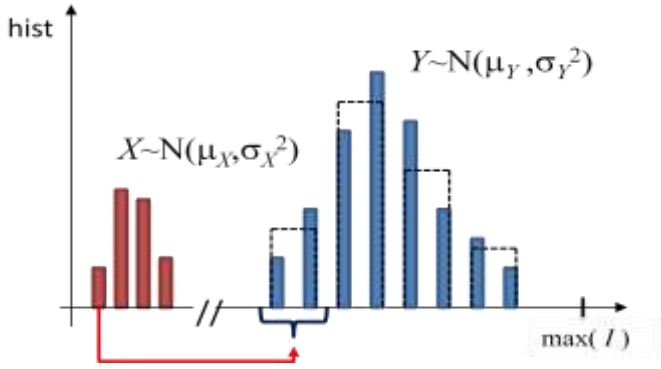


Figure 7.2. HM_{MGR} transformation example: for one bin of the shadow distribution, there exist two corresponding bins in the non-shadow distribution.

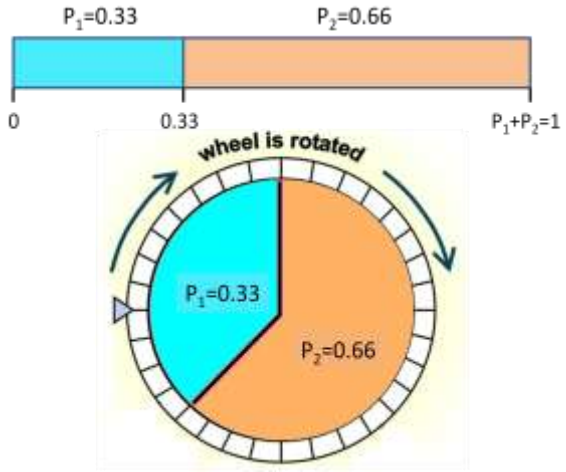


Figure 7.3. Roulette wheel for the two distributions in Figure 7.2.

works as a random sampling is used for this de-quantization by using the probabilities of y value. It is also called fitness proportionate selection [15]. In the illustration of Figure 7.2, it is supposed that the histogram bins ratio between X and Y is $R = 2$. The new non-shadow value will be selected by choosing randomly one of the $R = 2$ possible values, taking into consideration their occurrence (see Figure 7.3):

$$\hat{y} = sampling(P_1, \dots, P_R) \tag{7.4}$$

7.2.3. Gamma correction (GC)

The *GC* method is a standard method to encode luminance values in images [15]. This nonlinear function can also be adopted to restore shaded pixels x, as in [6]:

$$y = (x)^{\frac{1}{\gamma}}, \tag{7.5}$$

where γ is a constant value.

7.3. Reconstructability Problem Formulation

In VHR optical images, the presence of shadows may completely destroy the underlying information and thus can influence common processing and analysis procedures. They increase the risk to present false color tones, to distort the shape of objects, to merge, or to lose object. Shadow presences represent an important problem for both users and sellers of remote sensing images [6]. Information missing in shadow areas directly influences common processing and analysis operations, such as the generation of classification maps [17], [18]. An attempt solution is to restore as many shadow regions as possible in image I, by compensating the value of the shaded pixels by using one of the methods described in the previous section. Unfortunately, it may happen that not all shadow areas are possible to restore. Indeed, the main problem is the few radiometric levels characterizing the shadow areas compared to the corresponding non-shadow areas. Furthermore, shaded areas are usually characterized by a loss of texture [19].

To deal with this problem, before applying any shadow restoration method on I, we propose different criteria, capable to evaluate *a priori* if it is possible or not to reconstruct a given shaded area. To do that, human help is needed to recognize in I the thematic classes ($\bar{S}_i$) and the corresponding shadowed classes (S_i). Consequently, the histograms of the two classes are estimated for each thematic class. Given in input to a reconstruction criterion the two histograms, the criterion will aim at understanding if the corresponding shadow class is reconstructable or not (see illustration in Figure 7.4).

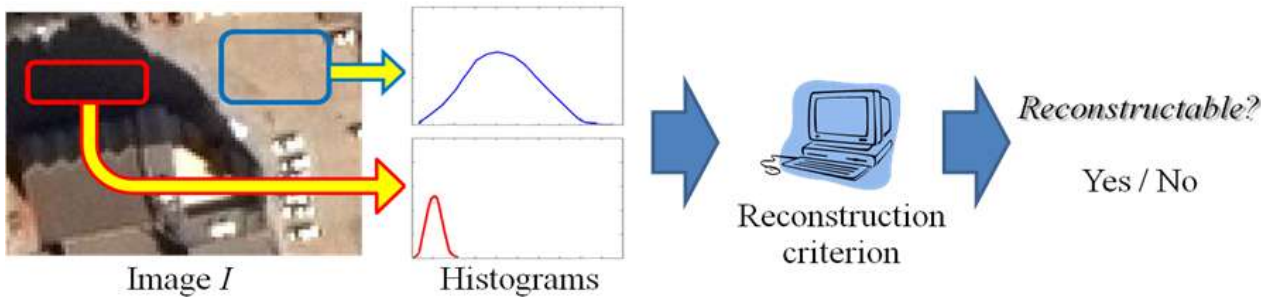


Figure 7.4. Key steps of the proposed shadow reconstructability analysis.

7.4. Proposed Criteria

In this section, eight evaluation criteria are detailed. Note that some of them exploit statistics estimated from the distribution of pixel intensities in image I (at pixel-level), while others are linked to higher-order statistics, estimated on the whole image, exploiting neighborhood information of pixels (at texture-level). The criteria are defined by comparing in some way the shadow and non-shadow histograms.

7.4.1. Histogram quantization error (HQE)

This criterion starts from the idea to consider the shadowing as a rescaling and quantization process [20]. As a first consequence, under this idea, we obtain two distributions which have different mean values. Therefore the first operation is to remove the mean value from both histograms in order to isolate the quantization component, i.e., $X' = X - \mu_X$ and $Y' = Y - \mu_Y$. In particular, we evaluate the histogram quantization error as follows:

$$\mathrm{Err}^2 = \sum_{i=1}^{\#bin(X')} (y'_i - x'_i)^2 P(x'_i)\,, \tag{7.6}$$

where $\#bin(X')$ corresponds to the number of bins in the histogram of the shadow class (note that $\#bin(X') = \#bin(X)$). This quantization error will be used as a potential criterion of shadow reconstructability evaluation:

$$C_{HQE} = \mathrm{Err}^2\,. \tag{7.7}$$

According to this formulation, the larger the criterion value, the larger the likelihood that X is unreconstructable.

7.4.2. Gray level ratio (GLR)

The *GLR* criterion evaluates the ratio between the supports of the two distributions. As in HM_{MGR}, here the evaluation index compares through a ratio operator the number of bins (different from zero) of the non-shadow distribution with the one of the shadow distribution:

$$C_{GLR} = \frac{\#\mathrm{bin}^*(Y)}{\#\mathrm{bin}^*(X)}, \tag{7.8}$$

where $\#\mathrm{bin}^*(\cdot)$ corresponds to the number of nonzero bins (i.e., those for which $P(x_i') \neq 0$ and $P(y_i') \neq 0$). Note that the larger C_{GLR}, the larger the probability that X is unreconstructable.

7.4.3. Two sample Kolmogorov-Smirnov test (2KS)

The *2KS* is a nonparametric test which compares one sample with a reference probability distribution. It can also be adapted to examine two distributions as in this case [20]. In particular, it tests whether two distributions, $f_X(y)$ and $f_Y(x)$, are coming from the same probability density function (null-hypothesis H_0) or not (alternative hypothesis H_1) [21]. In this case, the distance is evaluated by the maximum absolute difference over the support of the distribution $z \in [\min(I)\,, \max(I)]$:

$$D(z) = \max_z |\mathrm{CDF}_{X'}(z) - \mathrm{CDF}_{Y'}(z)|.$$

This distance will be considered as a potential criterion for our reconstruction problem:

$$C_{2KS} = D\,. \tag{7.9}$$

Regarding the reconstructability, C_{2KS} as well as the criteria defined in the following subsections, behave like C_{HQE} and C_{GLR}.

7.4.4. Variance Ratio (VR)

The variance is a common second-order statistical measure which evaluates how much a set of numbers are spread out from the mean value [22]. The considered evaluation criterion compares the variances of two histograms by rationing:

$$C_{VR} = \frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}. \tag{7.10}$$

7.4.5. Negentropy difference (ND)

In statistics, negentropy is adopted as a measure of distance to a normal distribution. Negentropy is always nonnegative and vanishes if and only if the signal is gaussian. According to [23], given a certain distribution X, negentropy is defined as:

$$NEG(X) = H(N_X) - H(X), \tag{7.11}$$

where $H(N_X)$ is the entropy of a normal distribution with the same covariance matrix as X, and $H(X)$ stands for the Shannon entropy of X. $NEG(Y)$ can be evaluated in the same way for the non-shadow area. In order to overcome the problem that the two distributions may have few bins, a kernel density estimation (KDE) is performed [24]:

$$f_h(x) = \frac{1}{n\,h}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right), \tag{7.12}$$

with $K(\cdot)$ being a gaussian kernel, n is the number of samples and h is a smoothing parameter useful to get a scaled kernel. This last parameter can be chosen as: $h = 3.491\sigma/n^{1/3}$ which will minimize the quadratic error of the pdf estimation when it is considered to follow a Gaussian law [24]. At the end, as an evaluation criterion, the following absolute difference function is considered

$$C_{ND} = |NEG(Y') - NEG(X')|. \tag{7.13}$$

7.4.6. Kullback-Leibler Divergence (DKL)

Other methods to compare the shape of two distributions are the f-divergences [25]. These methods measure the expectation of the diversity of the likelihood ratio between two distributions X and Y:

$$D_f(X|Y) = E_Y\left[f\left(\frac{dx(u)}{dy(u)}\right)\right] = \int f\left(\frac{x(u)}{y(u)}\right) y(u)du\,, \tag{7.14}$$

where E_Y is the expectation with respect to Y and f is a continuous and convex function. In the literature there exist several f-functions (among them Chi-2, Hellinger, …). A particular popular function is $f(u) = \log(u)u$ which leads to the so-called Kullback-Leibler divergence [26]. Substituting it in (7.14), the divergence measure is:

$$D_{KL}(Y|X) = \int \log\left(\frac{x(u)}{y(u)}\right) x(u)du. \tag{7.15}$$

Due to the fact that (7.15) is not a symmetrical measure, it is possible to rewrite it, obtaining a symmetrical formulation:

$$\frac{1}{D_{KL}(X,Y)} = \frac{1}{D_{KL}(X|Y)} + \frac{1}{D_{KL}(Y|X)}. \tag{7.16}$$

Note that also in this criterion the results come from distributions obtained with a gaussian KDE (kernel density estimation) method, applied on X' and Y'. As an evaluation index, we will explore:

$$C_{DKL} = D_{KL}(X', Y'). \tag{7.17}$$

7.4.7. Angular Second-Moment Difference (ASMD)

The two following criteria take into account second-order texture statistics, based on the gray level co-occurrence matrix GLCM, with dimension $M \times M$ [27]-[28]. The first is the angular second moment (ASM), defined as:

$$\text{ASM} = \sum_{i=1}^{M} \sum_{j=1}^{M} \big(P(i,j)\big)^2, \tag{7.18}$$

with (i,j)th entry in a gray-tone spatial dependence matrix $P(i,j)$ and M is the number of distinct gray levels in the image. The ASM measures the homogeneity, or in other words the textural uniformity in an image or a region of it. For example, if the considered window contains only similar gray levels, few elements of the GLCM convey nonzero values, whereas most of them are close to zero. Assuming the RV of shadow class as $X \sim \mathrm{N}\big(\mu_X, \sigma_X^2\big)$ and the corresponding RV of non-shadow class as $Y \sim \mathrm{N}\big(\mu_Y, \sigma_Y^2\big)$, here as evaluation criterion, we will exploit the following relative difference:

$$C_{ASMD} = \frac{|\text{ASM}_X - \text{ASM}_Y|}{\text{ASM}_Y}. \tag{7.19}$$

7.4.8. Homogeneity Difference (HD)

The second texture measure adopted in this article is the homogeneity (Hom), also called inverse difference moment. This measure assumes higher values for small differences of a pair of elements. This statistic is more sensitive to the presence of near diagonal elements in the GLCM. It can be evaluated as [27]-[28]:

$$\text{Hom} = \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{P(i,j)}{1+(i-j)^2}. \tag{7.20}$$

Assuming $X \sim \mathrm{N}\big(\mu_X, \sigma_X^2\big)$ and $Y \sim \mathrm{N}\big(\mu_Y, \sigma_Y^2\big)$, we will consider as last evaluation criterion the following relative difference:

$$C_{HD} = \frac{|\text{Hom}_X - \text{Hom}_Y|}{\text{Hom}_Y}. \tag{7.21}$$

7.5. Criterion Selection and Decision Making

7.5.1. Criterion selection

The goal of this subsection is to combine all previous criteria into a unique scalar value. Merging different information sources (in our case, different criteria) may be worth to obtain a

better answer to the faced problem (here, shadow reconstruction) [29]. A *t*-norm fuzzy logic combination is adopted in this work, where *t* stands for triangular and refers to the fact that in the framework of probabilistic metric spaces *t*-norms are adopted to generalize triangle inequality of ordinary metric spaces [30]. In particular, the minimum *t*-norm is used:

$$C_{FLC} = \perp_{min} (C_1, C_2, \dots, C_8) = \min\{C_1, C_2, \dots, C_8\}, \tag{7.22}$$

where C_i corresponds to one of the criteria presented above. Note that before merging the criteria, each of them is normalized in the range [0, 1] by means of the following relationship:

$$\tilde{C}_i = \frac{C_i}{1+|C_i|}. \tag{7.23}$$

From this combination, it is also intended to understand which criteria convey more information regarding the reconstructability of shadow areas and therefore retain the best ones. For such purpose, we will adopt a simple scheme that consists in proceeding by elimination, i.e., at each iteration remove from the list of criteria the poorest single criterion, up to get a subset of criteria which maximizes the accuracy.

7.5.2. Reconstruction evaluation

For a given criterion, the problem becomes now the one to find the threshold value (Th), which will act as the reference for the criterion on the basis of which a given shadow class will be assigned as a reconstructable or unreconstructable class. To estimate such a threshold, we will adopt the zero false alarm rate (ZFAR) principle in order to limit as most as possible reconstruction problems. In particular, it will not be tolerated that an unreconstructable area (UA) is assigned as a reconstructable area (RA) while the inverse is permitted. Under a ZFAR perspective, we will select the threshold value (Th) which maximizes the detection accuracy (ACC). In a more formal way, let C be one of the criteria defined above. Let $p(C|UA)$ and $p(C|RA)$ be the probability density function conditioned to unreconstructable and reconstructable areas (UA and RA), respectively. Since the above-defined criteria are such that the larger the criterion value, the larger the likelihood that the considered area is unreconstructable, Th is such that:

$$\begin{cases} C \leq Th \ \rightarrow \text{ area is reconstructable} \\ otherwise, \text{it is unreconstructable} \end{cases} \tag{7.24}$$

Accordingly, FAR is defined as:

$$\text{FAR}(Th) = \int_0^{Th} p\,(C|UA)\,dC, \tag{7.25}$$

while ACC is given by:

$$\text{ACC}(Th) = \int_0^{Th} p\,(C|RA)\,dC, \tag{7.26}$$

As illustrated in Figure 7.5, the best threshold value is such that:

$$Th^* = \underset{Th}{\text{argmax}}\{\text{ACC}(Th)\} \ \ s.t. \ \text{FAR}(Th) = 0. \tag{7.27}$$

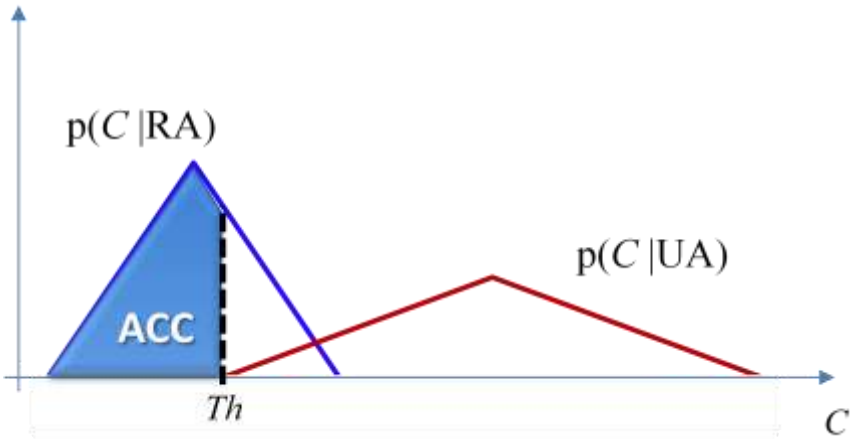


Figure 7.5. Illustration of the estimation of the best threshold value according to the ZFAR principle.

7.6. Experimental Results

7.6.1. Dataset description

Three different images have been considered. The first one refers to a 450×600 pixels QuickBird image crop (4 spectral bands, with a resolution of 0.6 meter), acquired on the 23rd of May 2003 over a rural region of Boumerdes in Algeria (see Figure 7.6(a)). The second and third images were acquired by IKONOS (only RGB spectral bands, with a resolution of 1 meter). The second one represents a crop of 450×600 pixels of the city of Riyadh (Saudi Arabia) acquired on the 28th of April 2008 (see Figure 7.6(b)), while the last one represents the center of the city of Atlanta (USA) with a crop of 420×500 pixels and was taken in 1998 (see Figure 7.6(c)). All these images contain different land covers and, for this reason, also different shadow classes: namely, shadows related to vegetation or bare soil areas, asphalt surface, sidewalk or roofs, etc. (see second column in Table 7.I). Note that we exploit all the spectral bands in our experiments.

7.6.2. Experimental setup

In order to assess the reconstruction criteria described above, we have to quantify in some way the concept of reconstructability of shadow classes. For such purpose, human help is required to decide if an area has been reconstructed satisfactorily or not and thus is reconstructable or not. In more detail, first two areas are chosen to define each thematic class and the corresponding shadow class present in each image. This step is iterated 3 times, to produce various regions for each class. At the end, for the 12 thematic classes characterizing the three images, we have produced $12 \times 3 = 36$ regions. Each shadow class has been reconstructed by adopting the above-described reconstruction methods. In total, we got 5 reconstruction methods, multiplied by 36 regions, providing thus 180 possible reconstructions. Afterwards, a human expert evaluated visually each

(a)

(b)

(c)

Figure 7.6. Images of (a) Boumerdes, (b) Riyadh, and (c) Atlanta datasets.

reconstruction, in order to infer a decision about the reconstructability of the shadow classes. In particular, due to the fact that we have for the same shadow class, three different regions and five reconstruction methods, the human expert returned a positive judgment ("YES") only if all the 15 reconstructed shadow regions are well restored; negative ("NO") otherwise. Due to the fact that the *GC* method failed to return a satisfactory reconstruction for all the 12 thematic classes, it was

TABLE 7.I
RECONSTRUCTION RESULTS FOR THE DIFFERENT CLASSES.

Dataset	Class	Reconstruction Methods						Reconstruct-ability
		LT	*HM*			*GC*	Sum/Max	
			rank	*MGR*	*RGR*			
Boumerdes	1. vegetation	3	3	3	3	0	12/12	YES
	2. asphalt	3	0	3	3	0	9/12	NO
	3. bare soil	3	0	0	0	0	3/12	NO
	4. sidewalk	3	3	3	3	0	12/12	YES
Riyadh	1. vegetation	3	3	3	3	0	12/12	YES
	2. asphalt	2	0	0	0	0	2/12	NO
	3. roofs	3	3	3	3	0	12/12	YES
	4. tiles	0	0	0	0	0	0/12	NO
	5. parking	3	3	3	3	0	12/12	YES
Atlanta	1. bright roofs	3	3	3	3	0	12/12	YES
	2. asphalt	0	0	0	0	0	0/12	NO
	3. vegetation	3	3	3	3	0	12/12	YES

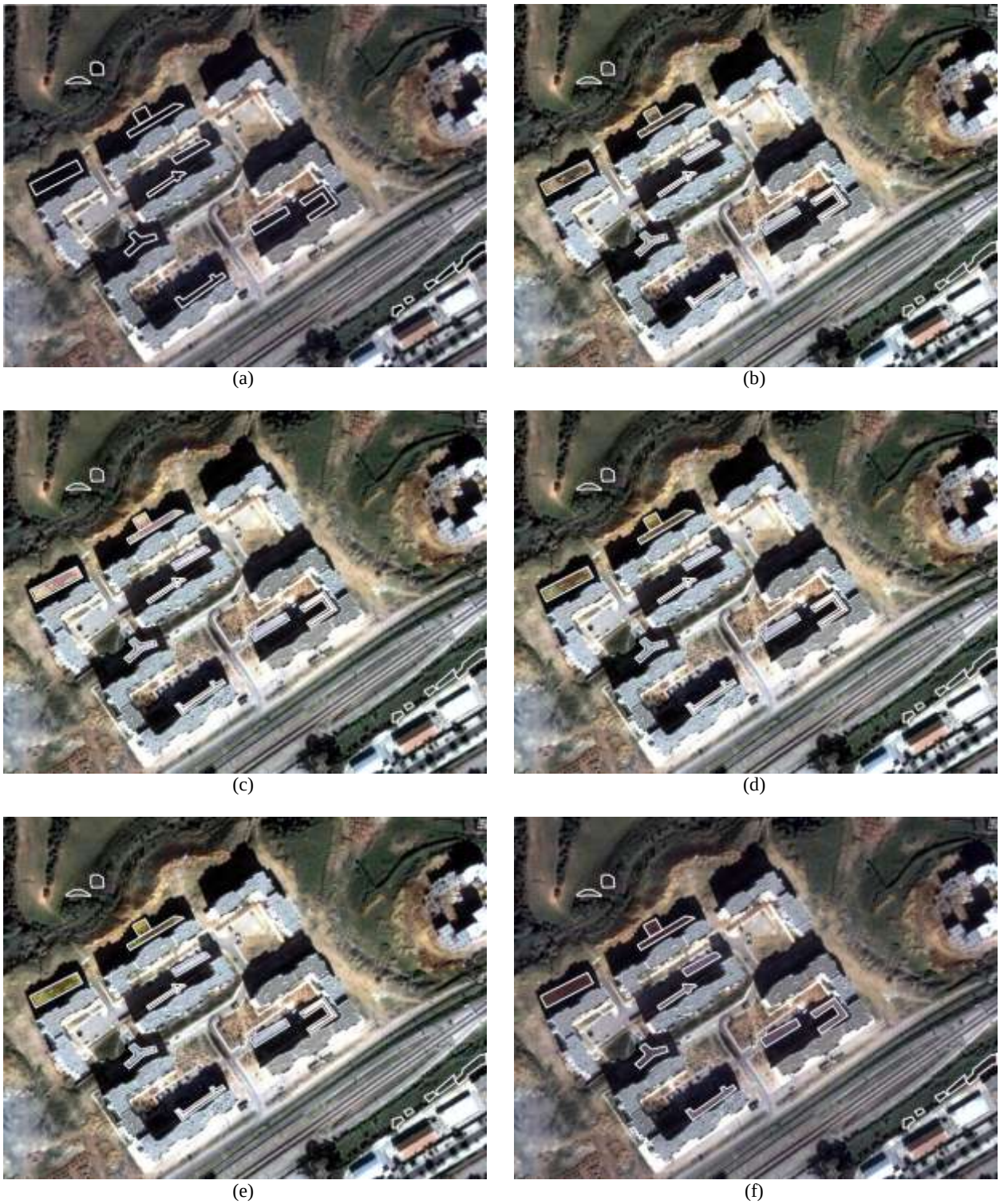
discarded from the pool of reconstruction methods (see Table 7.I). The unsatisfactory results of *GC* can be explained by the absence of bias compensation in its formulation.

From Table 7.I, we can observe that the *LT* method returns the largest number of positive votes compared to the other methods. Furthermore, all the methods (except the excluded *GC*) return almost similar results which simplify the final decision about the reconstructability or not of each shadow class. As mentioned above, we adopted the following rule: "a shadow class is reconstructable if all four reconstruction method succeed in its satisfactory reconstruction". Figure 7.7 shows for illustration the reconstruction of some regions of the shadow classes extracted from the Boumerdes image. In general, it is possible to deduce that class 1 (vegetation) and class 4 (sidewalk) are well restored, whereas class 2 (asphalt) and class 3 (bare soil) present poor reconstructions, with in some cases unacceptable visual artifacts.

7.6.3. Results

The present experiments have the scope to compare the several criteria defined above. Starting from the final result obtained before (the last column of Table 7.I), we first merged for each image the three regions representing the same class in order to compute the eight proposed criteria for each thematic class in each image. Table 7.II shows the values obtained for each evaluation criterion and for each of the twelve classes. We recall that the first six criteria work at pixel-level, whereas the next two at the texture-level. These results were useful to compute, for each criterion, the best threshold value (first raw of Table 7.III), obtained by requiring a zero value for the false alarm probability ($FAR = 0$) and maximizing the reconstruction accuracy (ACC), as formulated in (7.26).

From the graphs in Figure 7.8, we can see for each criterion the trend of the FAR and ACC by varying the threshold value (Th). The numerical results satisfying (7.27) are reported in Table 7.III. In particular, when a zero value is required for FAR, and ACC is maximized, it comes out that the



(a) (b) (c) (d) (e) (f)

Figure 7.7. (a) Original Boumerdes image with shadow areas highlighted. Shadow areas reconstructed with: (b) *LT*, (c) HM_{rank}, (d) HM_{MGR}, (e) HM_{RGR} and (f) *GC*.

best criterion is the angular second-moment difference (C_{ASMD}). This criterion return an accuracy $ACC = 0.71$, with a zero false alarm rate ($FAR = 0$).

We then fused all the criteria to yield a unique consensus criterion following the *t*-norm fuzzy combination defined in (7.22). As explained above, the min operator is used in such a fusion. Combining all the eight criteria yields a poor result, i.e., accuracy $ACC = 0.14$ and zero false alarm

TABLE 7.II
NUMERICAL RESULTS OBTAINED FROM THE INVESTIGATED NORMALIZED EVALUATION CRITERIA.

Dataset	Class	From Table 6.I	C_{HQE}	C_{GLR}	C_{2KS}	C_{VR}	C_{ND}	C_{DKL}	C_{ASMD}	C_{HD}
Boumerdes	1. vegetation	Yes	0.333	0.431	0.027	0.390	0.073	0.266	0.171	0.158
	2. asphalt	No	0.517	0.703	0.044	0.836	0.021	0.174	0.469	0.183
	3. bare soil	No	0.865	0.871	0.127	0.994	0.064	0.301	0.711	0.474
	4. sidewalk	Yes	0.897	0.849	0.180	0.980	0.020	0.259	0.701	0.791
Riyadh	1. vegetation	Yes	0.837	0.720	0.095	0.867	0.052	0.133	0.616	0.330
	2. asphalt	No	0.847	0.774	0.120	0.925	0.072	0.284	0.729	0.546
	3. roofs	Yes	0.623	0.525	0.039	0.472	0.066	0.114	0.087	0.218
	4. tiles	No	0.908	0.795	0.192	0.957	0.046	0.181	0.509	0.747
	5. parking	Yes	0.934	0.792	0.178	0.956	0.031	0.125	0.099	0.158
Atlanta	1. bright roofs	Yes	0.329	0.460	0.070	0.427	0.044	0.232	0.025	0.180
	2. asphalt	No	0.875	0.663	0.242	0.906	0.182	0.275	0.238	0.511
	3. vegetation	Yes	0.899	0.624	0.145	0.674	0.033	0.283	0.032	0.150

TABLE 7.III
BEST THRESHOLD VALUE OBTAINED FOR EACH CRITERION AND CORRESPONDING ACC VALUE.

	C_{HQE}	C_{GLR}	C_{2KS}	C_{VR}	C_{ND}	C_{DKL}	C_{ASMD}	C_{HD}
Th*	0.517	0.663	0.044	0.836	0.021	0.174	0.238	0.183
ACC	0.29	0.57	0.29	0.57	0.14	0.43	**0.71**	0.57

TABLE 7.IV
RESULTS OF THE CRITERION SELECTION PROCEDURE BASED ON ITERATIVE ELIMINATION.

Fusion Rule	# Criteria	ACC	*Th**
$\min\{C_{HQE}, C_{GLR}, C_{2KS}, C_{VR}, C_{ND}, C_{DKL}, C_{ASMD}, C_{HD}\}$	8	0.14	0.021
$\min\{C_{HQE}, C_{GLR}, C_{2KS}, C_{VR}, C_{DKL}, C_{ASMD}, C_{HD}\}$	7	0.57	0.044
$\min\{C_{HQE}, C_{GLR}, C_{VR}, C_{DKL}, C_{ASMD}, C_{HD}\}$	6	0.86	0.174
$\min\{C_{GLR}, C_{VR}, C_{DKL}, C_{ASMD}, C_{HD}\}$	5	0.86	0.174
$\min\{C_{VR}, C_{DKL}, C_{ASMD}, C_{HD}\}$	4	0.86	0.174
$\min\{C_{DKL}, C_{ASMD}, C_{HD}\}$	3	0.86	0.174
$\min\{C_{DKL}, C_{ASMD}\}$	**2**	**0.86**	0.174

(see first row of Table 7.IV). This is explained by the presence of poor criteria in the pool. To try to obtain better results, we iteratively eliminated the worst criteria from the pool. Accordingly, first, we eliminated the negentropy criterion (C_{ND}). The accuracy improved to $ACC = 0.57$ (see Table 7.IV). Afterwards, we removed the Kolmogorov-Smirnov criterion (C_{2KS}). This allowed to raise further the accuracy to $ACC = 0.86$, which is better than what is achieved by the best single criterion (C_{ASMD}). We pushed further the elimination process to reduce the pool to just two criteria. The accuracy kept unchanged but with the advantage of simplifying the pool to C_{DKL} and C_{ASMD}. This means that, with a threshold value of $Th^* = 0.174$ there is no risk that a reconstructable shadow area will be detected as reconstructable and 86% of the reconstructable shadow areas are correctly detected. From this last result, it emerges that in order to restore a shadow area both histogram and texture comparison between shadow and non-shadow regions of the same class have

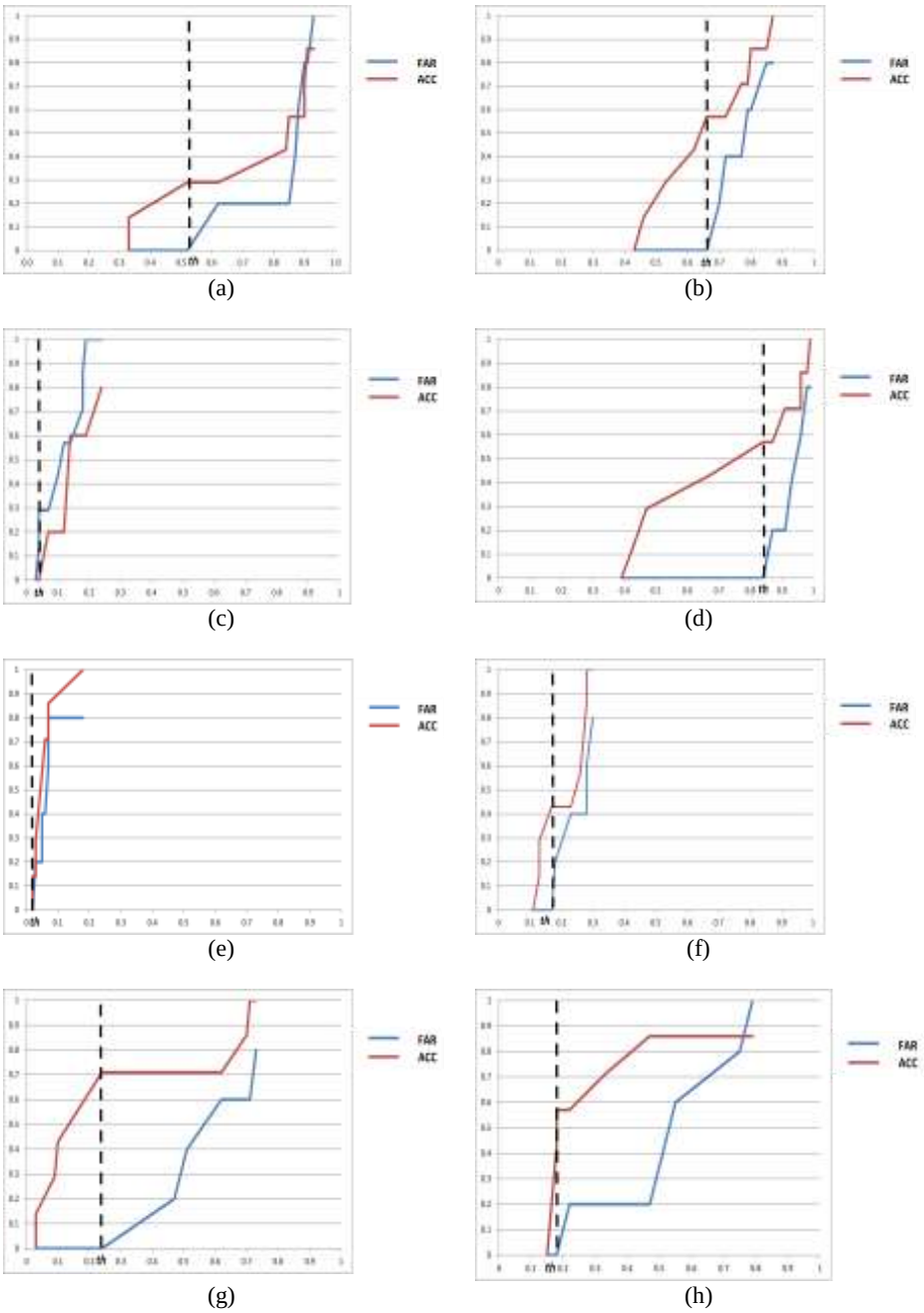


Figure 7.8. Behavior of the false alarm rate (FAR) and the accuracy (ACC) values for the eight explored evaluation criteria (normalized between 0 and 1). (a) C_{HQE}, (b) C_{GLR}, (c) C_{2KS}, (d) C_{VR}, (e) C_{ND}, (f) C_{DKL}, (g) C_{ASMD}, (h) C_{HD}.

to be taken into account. To synthesize, all previous empirical results lead us to propose the following Shadow Area Recovering Index:

$$\text{SARI} = \min\{C_{DKL}, C_{ASMD}\}. \tag{7.23}$$

7.6.4. Validation of the Results

To validate our methodology and the proposed SARI index, we consider three new crops coming from the three original images. These new crops may contain similar but possibly different thematic classes (see Figure 7.9). We estimate on such images the two criteria (C_{DKL} and C_{ASMD}) and from these last results we evaluate the minimum value: If SARI returns a value greater than the previous computed threshold value ($Th^* = 0.174$), this means that it is not possible to obtain a good reconstruction of such shadow class. Vice versa, if SARI value results below the threshold value, we are sure that it is possible to restore such shadow class. Table 7.V shows the numerical results obtained by the two criteria (C_{DKL} and C_{ASMD}) and their minimization via the SARI criterion. Figure 7.10 illustrates the original shadow classes coming from the test images, and only the reconstructable classes coming from Table 7.V, this means, only when SARI criterion returns a value lower than the threshold. Note that, to perform the reconstructions, in this case, we simply adopt the linear transformation (LT) method. All these well-done reconstructions represent a positive judgment for the proposed SARI index.

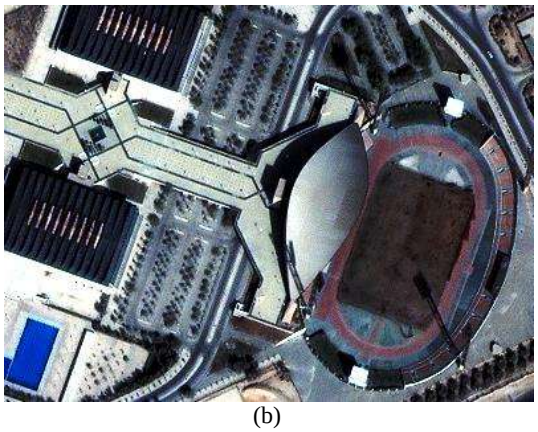
TABLE 7.V
NUMERICAL SARI RESULTS OBTAINED ON THE TEST IMAGES AND THEIR RECONSTRUCTABILITY JUDGMENT. NOTE THAT $Th^* = 0.174$.

		C_{DKL}	C_{ASMD}	SARI	Reconstruct-ability
Boumerdes 2	1. vegetation	0.133	0.206	0.133	Yes
	2. asphalt	0.119	0.245	0.119	Yes
	3. bare soil	0.213	0.113	0.113	Yes
Riyadh 2	1. vegetation	0.224	0.216	0.216	No
	2. tiles	0.131	0.233	0.131	Yes
	3. asphalt	0.214	0.266	0.214	No
	4. dark roofs	0.247	0.326	0.247	No
Atlanta 2	1. bright roofs	0.108	0.142	0.108	Yes
	2. asphalt	0.227	0.216	0.216	No
	3. bare soil	0.231	0.113	0.113	Yes
	4. vegetation	0.131	0.099	0.099	Yes

7.7. Conclusion

This chapter deals with the important problem to determine if it is possible to infer *a priori* if a determinate shadow class is restorable. A set of different criteria has been proposed. Starting from two histograms, one of the shadow class and the other of the corresponding non-shadow class, they aim at providing an objective assessment of the shadow reconstructability. They differ from each other depending on their complexity level, the statistical information they need, and if they exploit or not texture information.

(a)



(b)

(c)

Figure 7.9. Test images of (a) Boumerdes 2, (b) Riyadh 2, and (c) Atlanta 2 datasets.

In the experiments which are based on three different images showing various shadow contamination scenarios, we have seen that the explored reconstruction methods return similar reconstruction results, except the *GC* method. The experiments show also that not all the evaluation criteria return satisfactory results. The poorest criteria are those based on the histogram quantization error, the two sample Kolmogorov-Smirnov test and the negentropy difference, suggesting that full

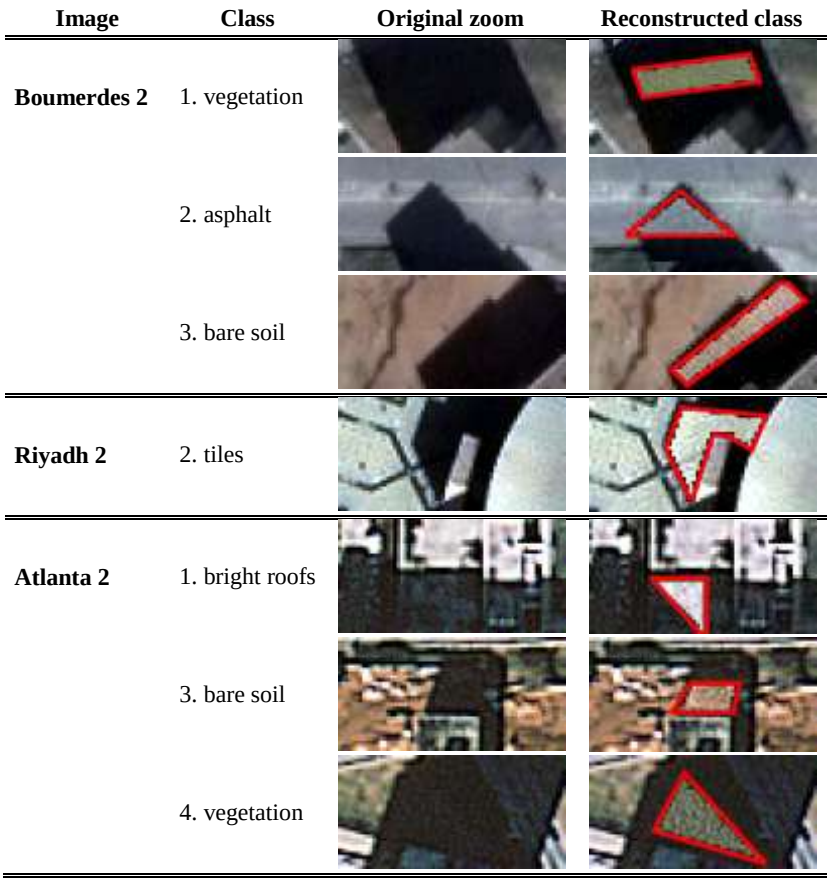


Figure 7.10. Original and good reconstruction of the shadow classes from the test images, adopting the linear transformation (*LT*) method.

histogram comparison may be misleading due to the difficulty to get a reliable between-histogram distance measure in particular when one of the histograms (in this case, the shadow histogram) is characterized by a small number of bins. By contrast, criteria such as gray level and variance ratios, and angular second-moment difference behave better as they compare directly or indirectly the number of bins conveyed by the two histograms or they take advantage from the textural domain. The fusion of criteria proved to be a useful way, though more computationally demanding, to capture the synergies between the different criteria. In particular, it allows us suggesting as a final criterion, the shadow area recovery index (*SARI*), limited to the estimation of the Kullback-Leibler divergence and the angular second-moment difference.

As an additional future work, it will be interesting to substitute the need of a human help, to decide if an area has been reconstructed satisfactorily or not, with an automatic analysis.

7.8. Acknowledgment

The authors are grateful to CNES for making available the IKONOS-2 Atlanta image through the Orfeo initiative.

7.9. References

[1] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 442–455, Feb. 2006.

[2] S. Benabdelkader and F. Melgani, "Contextual spatiospectral postreconstruction of cloud-contaminated images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 204–208, Apr. 2008

[3] B. Abdel Latif, R. Lecerf, G. Mercier, and L. Hubert-Moy, "Preprocessing of low-resolution time series contaminated by clouds and shadows," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2083–2096, Jul. 2008.

[4] L. Lorenzi, F. Melgani, and G. Mercier, "Inpainting strategies for reconstruction of missing data in VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 914–918, Sep. 2011.

[5] P. Sarabandi, F. Yamazaki, M. Matsuoka, and A. Kiremidjian, "Shadow detection and radiometric restoration in satellite high resolution images," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 6, pp. 3744–3747.

[6] A. Massalabi, H. Dong-Chen, G.B. Benie and E. Beaudry, "Detecting information under and from shadow in panchromatic IKONOS images of the city of Sherbrooke," in *Proc. IGARSS*, Anchorage, AK, USA, Sep. 2004, vol. 3, pp. 2000–2004.

[7] J. Su, X. Lin, and D. Liu, "An automatic shadow detection and compensation method for remote sensed color images," in *Proc. ICSP*, Beijing, China, Nov. 2006, vol. 2.

[8] T. Nakajima, G. Tao, and Y. Yasuoka, "Simulated recovery of information in shadow areas on IKONOS image by combining ALS data," in *Proc. ACRS*, Kathmandu, Nepal, 2002.

[9] F. Yamazaki, W. Liu, and M. Takasaki, "Characteristic of shadow and removal of its effects for remote sensing imagery," in *Proc. IGARSS*, Cape Town, South Africa, Jul. 2009, vol. 4, pp.426–429.

[10] L. Lorenzi, F. Melgani, and G. Mercier, "A complete processing chain for shadow detection and reconstruction in VHR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3440–3452, Sep. 2012.

[11] A. Makarau, R. Richter, R. Müller, and P. Reinartz, "Adaptive shadow detection using a blackbody radiator model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2049–2059, Jun. 2011.

[12] V. Achard, X. Briottet and L. Poutier, "Scope of ONERA atmospheric compensation tools to retrieve the optical properties over different type of landscape from hyperspectral airborne acquisition in the [0.4 – 2.5 um] domain," in *Proc OPTRO*, Paris, France, Feb. 2010.

[13] D. Shen, "Image registration by local histogram matching," *Pattern Recognition Journal*, vol. 40, no. 4, pp. 1161–1172, Apr. 2007.

[14] H.A. David and H.N. Nagaraja, *Order statistics*, John Wiley & Sons, 3rd edition, 2003.

[15] J. Holland, *Adaptation in natural and artificial systems*, Bradford Book, 1992.

[16] W. Burger and M. Burge, *Principles of digital image processing: Fundamental techniques*, Springer, 2006.

[17] T. Kasetkasem and P.K. Varshney, "An optimum land cover mapping algorithm in the presence of shadow," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, Jun. 2011.

[18] B. Abdel Latif, R. Lecerf, G. Mercier, and L. Hubert-Moy, "Preprocessing of low-resolution time series contaminated by clouds and shadows," *IEEE Trans. Geosci. and Remote Sens.*, vol. 46, no. 7, pp. 2083-2096, Jul. 2008.

[19] P.M. Dare, "Shadow analysis in high-resolution imagery of urban areas," *J. of the American Society for Photogrammetry and Remote Sensing*, vol. 71, no. 2, pp. 169–177, Feb. 2005.

[20] B. Kamgar-Parsi, "Evaluation of quantization error in computer vision," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 11, no. 9, pp. 929–940, Sep. 1989.

[21] R. H. C: Lopez, I. Reid and P. R. Hobson, "The two-dimensional Kolmogorov-Smirnov test," in *Proc of Science*, Amsterdam, Netherlands, Apr. 2007.

[22] W. Navidi, *Statistics for engineers and scientists*, McGraw-Hill, 2006.

[23] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, vol. 13, pp. 411–430, 2000.

[24] A. W. Bowman and A. Azzalini, *Applied smoothing techniques for data analysis*, Oxford University Press, 1997.

[25] I. Csiszar, "Information type measures of difference of probability distributions and indirect observations," *Studia Sci. Math. Hungar.*, vol. 2, pp. 299–318, 1967.

[26] J. Inglada and G. Mercier, "A new statistical and 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.

[27] R. M. Haralick, K. Shanmugam and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. On System, Man and Cybernetics*, vol. 3, pp. 610–621, Nov. 1973.

[28] F. Pacifici, M. Chini and W. J. Emery, "A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification," *Journal of Remote Sensing of Environment*, vol. 113, no. 6, pp. 1276–1292, Mar. 2009.

[29] L. A. Klein, *Sensor and data fusion: a tool for information assessment and decision making*, SPIE Press, 2004.

[30] E.P. Klement, R. Mesiar and E. Pap, *Triangular norms*, Springer, 2000.

8. Conclusion

Abstract *– This chapter draws the conclusion of the research activity described in the thesis. In particular, it summarizes and discusses the methodological and experimental developments of each of the studies and presents an outlook of the possible future developments. The reader is referred to the previous single chapters for more detailed discussions about the different proposed methods.*

In this thesis, missing data problems on very high spatial resolution (VHR) optical satellite have been investigated, in order to detect and/or reconstruct the obscured areas. In particular, we face the problem of a complete obscuration given by the presence of cloudy areas, or partially obscured by the presence of shadow regions. We proposed several methodological aspects and solutions followed by a validation step based on real VHR remote sensing images. In the following, we will briefly summarize the conclusions drawn from each of the presented strategies. We refer the reader to the single chapters for more details.

In Chapter 3, *Inpainting Strategies for Reconstruction of Missing Data in VHR Images*, the problem of missing area reconstruction due to the presence of clouds was presented. We adopted several inpainting techniques on a single multispectral image, and in particular, three different strategies were described: feature extraction-based inpainting (FEBI), inpainting with isometric transformation (IsoI) and multiresolution inpainting (MRI). The experimental results obtained on real images with simulated clouds, show good capabilities of the proposed methods in the reconstruction of cloudy areas. In general, all the proposed methods were characterized by higher performance in terms of accuracies with respect to the region-based inpainting (RBI). Only the IsoI method did not return good accuracy results in almost all cases. In comparison with the rest, the strategy based on the addition of textural features, especially exploiting the Symlet feature, returned good results with respect to the reference method. However, the best performance in terms of PSNR was reached by the method based on the multiresolution inpainting, because of its progressive and multiresolution approach. In order to further improve the reconstruction accuracy, the inpainting techniques would need the integration of the temporal dimension.

In Chapter 4, *Support Vector Regression with Kernel Combination for Missing Data Reconstruction*, we proposed a new method specifically developed for support vector regression (SVR) of remote sensing multispectral images, in order to reconstruct missing data due to the presence of clouds. It integrates in the reconstruction process two types of information: 1) the position of the missing value and 2) the radiometric information. The best resulting combination of kernels was the adoption of two radial basis functions (RBFs). Their fusion performed by means of a kernel combination together with the power of the support vector regression made it particularly promising as suggested by the experiments. In more detail, we compared our methodology (GKCR) with four other reconstruction techniques, namely, MRI, CMLP, CNP and CSSPR. While, the first two techniques (MRI and CMLP) returned satisfactorily results, better performances were achieved by the others methods (CNP, CSSPR and GKCR) and also by varying the number of the training set. We also noticed that, with respect to the last two techniques and GKCR, GKCR has a greater accuracy and a lower complexity, namely a smaller number of support vectors. As a drawback, one has to pay a higher price in terms of computational time, especially for the tuning of the larger number of free parameters. In order to prevent this negative aspect, other techniques could be considered, such as quantile regression in a probabilistic space, which aims at estimating a quantile of the response variable.

In Chapter 5, *Missing Area Reconstruction in Multispectral Images Under a Compressive Sensing Perspective*, three new methods were presented, which, given a multitemporal image sequence with a contaminated image, allow the reconstruction of the obscured parts by exploiting compressive sensing (CS) theory. In particular, we first adopt two of the most common CS

solutions, capable to approximate the theoretical CS result, which are the basis pursuit (BP) and the orthogonal matching pursuit (OMP). In the second instance, we proposed to solve the theoretical CS through exploiting the capabilities of a genetic algorithm (GA). The experimental results obtained on a real dataset with a simulated presence of clouds, showed good capabilities of the proposed methods in the reconstruction of cloudy areas. In general, all the three proposed methods were characterized by good performance in terms of PSNR. In particular, OMP has the advantage of being sparser and significantly faster when compared with the other two methods, but at the same time less robust; BP, on the contrary, is the least sparse when compared with the other two. A good compromise in terms of sparseness between OMP and BP is represented by the GA solution, but its major drawback is the longer computational time. From the experimental results we obtained, it seems that the kind of ground cover obscured may be important, while the size of the contaminated area only marginally affects the performance of the proposed reconstruction methods. In order to improve the obtained results, the addition of textural features may help the methods in the reconstruction of missing data.

As regards the second part of the manuscript, in Chapter 6, *A Complete Processing Chain for Shadow Detection and Reconstruction in VHR Images*, a complete and supervised processing chain for the detection and removal of shaded regions in VHR images was proposed. Several image processing and pattern recognition tools were exploited, including first a detector for the shadow regions and then a support vector machine (SVM) classification in order identify the thematic classes and their shadows. Afterwards, we performed a specific reconstruction for each shadows class detected by the SVM. In particular, we performed a linear regression procedure, capable of adjusting the intensities of the shaded pixels according to the corresponding non-shadow regions. The whole methodology is capable of returning a shadow-free image preserving spectral and textural properties of the previous obscured regions, while other tested methods cannot. In particular, our method has the potential to check misreconstruction problems and handles the penumbra issue in a different way. Despite the interesting results obtained, some improvements are still necessarily, e.g., an *a priori* evaluation of the reconstructability of a shadow class, an active learning process to choose the most appropriate training set, a non-linear regressor for the cases where a linear regression does not work suitably, and a way to obtain in addition to a reconstruction map, a confidence map.

Finally, in Chapter 7, *Assessing the Reconstructability of Shadow Areas in VHR Images*, we have tried to address one of the future works from the previous chapter; although in general it can also be viewed as a common necessity for any shadow reconstruction method. An ideal reconstructability criterion should not tolerate that an unreconstructable shadow area is assigned as reconstructable and, at the same time, should maximize the probability of detection of reconstructable areas. Accordingly, the problem can be stated by the following question: is it possible to know *a priori* if a shadow area can be well recovered? Starting from two histograms, several evaluation criteria working at the pixel and textural level were proposed. They differ from each other depending on their complexity level, the statistical information they need, and if they exploit texture information or not. From experiments based on images acquired from three different sensor, with different kinds of shadow obscuration, it is shown that just part of the investigated evaluation criteria return satisfactory results. In particular, criteria such as gray level and variance

ratios, and angular second-moment behave better as they compare directly or indirectly the number of bins conveyed by the two histograms or they take advantage of the textural domain. The fusion of criteria allows us to suggest as a final criterion, the shadow area recovery index (SARI), which is limited to the estimation of the Kullback-Leibler divergence and the angular second-moment difference.

To conclude, the contributions provided in this thesis have been focused on the development of reconstruction techniques for missing data problem in very high resolution optical images, in particular for cloud and shadow areas. Such contributions have been critically analyzed considering the state-of-the-art of the related topics, and have been compared with reference methods by means of in-depth testing experiments. The obtained image reconstructions confirm that the research reported in this thesis has made interesting and original contributions to the faced methodological issues.

Since three years of research are never enough, future works about the problem of missing data can understandably be envisioned. For example, regarding the inpainting approach, it could be interesting to integrate the temporal dimension in the process of reconstruction. About the compressive sensing approach, at the moment, we have built-up a dictionary by adopting a trivial subsample methodology. We think that it could be interesting to opt for smart techniques to design the dictionary, focusing on the context of the missing area (e.g., for instance by local dictionary). Another possible development is the addition of features in the training step of the CS (e.g., Haralick textures, Hu invariant moments, …). Regarding the shadow area reconstruction, we may envision to resort to nonlinear method in case linear regression does not work in a proper way. Another idea is the formulation of the reconstruction problem within a Markov random field (MRF) framework. More in general, for each of the work presented here, it would be particularly interesting to reformulate or adapt it as view of an application for large-scale images (at continental or world levels), knowing that such images raise the problem of spatial variance. For such purpose, transfer learning techniques would be an interesting source of inspiration. Furthermore, we want to stress the reader that all the methods here proposed can be also applied not only in the case of missing data, but also in the case where we have available erroneous data.

9. List of Related Publications

9.1. Published Journal Papers

[J.1] L. Lorenzi, F. Melgani and G. Mercier, "Inpainting strategies for reconstruction of missing data in VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 6, pp. 914–918, Sep. 2011.

[J.2] L. Lorenzi, F. Melgani and G. Mercier, "A complete chain for shadow detection and reconstruction in VHR images," IEEE Trans. Geosci. Remote Sens., vol. 50, no. 9, pp. 3440–3452, Sep. 2012.

[J.3] L. Lorenzi, F. Melgani and G. Mercier, "A support vector regression with kernel combination for missing data reconstruction," *IEEE Geosci. Remote Sens. Lett.*, in press.

[J.4] L. Lorenzi, F. Melgani, G. Mercier and Y. Bazi, "Assessing the reconstructability of shadow areas in VHR images," *IEEE Trans. Geosci. Remote Sens*., in press.

9.2. Journal Papers in Revision

[J.5] L. Lorenzi, F. Melgani and G. Mercier, "Missing area reconstruction in multispectral images under a compressive sensing perspective," submitted to *IEEE Trans. Geosci. Remote Sens.*.

9.3. Conference Proceedings

[C.1] L. Lorenzi, F. Melgani and G. Mercier, "Multiresolution inpainting for reconstruction of missing data in VHR images," in *Proc. IGARSS*, Vancouver, Canada, Jul. 2011.

[C.2] L. Lorenzi, F. Melgani and G. Mercier, "Orthogonal matching pursuit for VHR image reconstruction," in *Proc. IGARSS*, Munich, Germany, Jul. 2012.

[C.3] L. Lorenzi, F. Melgani and G. Mercier, "Some criteria to assess the reconstructability of shadow areas," in *Proc. IGARSS*, Munich, Germany, Jul. 2012.