Change detection from photographs

Yan Wang

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *13/07/2016* **par :**

Yan WANG

Détection des changements à partir de photographies

Change Detection from Photographs

JURY

Philippe JOLY	Professeur, Université Toulouse 3	Président
Yan MA	Full Professor, Shanghai Normal University	Rapporteur
Laurent MASCARILLA	Maître de Conférences HDR, Université de La Rochelle	Rapporteur
Alain CROUZIL	Maître de Conférences, Université Toulouse 3	Directeur de thèse
Jean-Baptiste PUEL	Maître de Conférences, ENFA Toulouse-Auzeville	Co-directeur de thèse

---

École doctorale et spécialité :
MITT : Image, Information, Hypermédia
Unité de Recherche :
Institut de Recherche en Informatique de Toulouse (UMR 5505)
Directeur(s) de Thèse :
Alain CROUZIL et *Jean-Baptiste PUEL*
Rapporteurs :
Yan MA et *Laurent MASCARILLA*

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Alain Crouzil and my co-supervisor Jean-Baptiste Puel, for their guidance, support, valuable discussions and thorough comments. Three years ago, when I left my home country for the first time and started the PhD studies in France, I always get into trouble due to the different language, lives and the difficulties occurring in my research work. My supervisors gave me great encouragements and support. Their endless patience made me feel warm and made me insist on my research work although it is so difficult. With their guidance and insightful comments, I can gradually overcome obstacles and enjoy my work. None of this dissertation would be possible without their invaluable aids.

I would like to warmly thank my jury members, Prof. Philippe Joly, Prof. Yan Ma and Dr. Laurent Mascarilla, for serving as my jury members. I have benefited a lot from their insightful reviews, comments and questions about my work.

My warmest thanks should go to my lab mates: Remi Dubot, Arturo Curiel and Bérengère Mathieu for all the pleasing discussions, help, and encouragements in my study and my lives during the last three and half years. I sincerely thank all the faculties in ENFA. Special thanks should be given to Jean-Luc Granier, Lucile Plenecassagnes, Loic Barreau. Although I can not contact with them in French, everyday, their smiles and laugh made me feel at home.

Many thanks also go to all of my friends: Xuan Liu, Tiantian Xiong, Rui Sang, Ran Zhao, Xiaobo Wang, Dan Wang, Wenjun Zhang and Meizhi Huang for providing me great fun and help.

I am grateful to Chinese Scholarship Council for providing funding for my study in France.

Last, but not the least, I would like to thank my husband and parents for their love and supporting me spiritually throughout my life. My husband, Youlong WU, is always behind me and gives his unconditional support. During my PhD studies, he always helps me whenever I face difficulties and encourages me whenever I feel upset. For each paper or report I wrote, he always reads carefully and gives me insightful comments and advices. Thanks to my parents for raising me and giving me the best education they can offer.

For these three persons, I know that I can never fully repay them, but I will spend the rest of my life trying.

Dedication

To my family.

Abstract

This work deals with change detection from chronological series of photographs acquired from the ground.

This context of consecutive images comparison is the one encountered in the field of integrated geography where photographic landscape observatories are widely used. These tools for analysis and decision-making consist of databases of photographic images obtained by strictly rephotographing the same scene at regular time intervals. With a large number of images, the human analysis is tedious and inaccurate. So a tool for automatically comparing pairs of landscape photographs in order to highlight changes would be a great help for exploiting photographic landscape observatories. Obviously, lighting variations, seasonality, time of day induce completely different images at the pixel level. Our goal is to design a system which would be robust to these insignificant changes and able to detect relevant changes of the scene.

Numerous studies have been conducted on change detection from satellite images. But the utilization of classic digital cameras from the ground raise some specific problems like the limitation of the spectral band number and the strong variation of the depth in a same image which induces various appearance of the same object categories depending on their position in the scene.

In the first part of our work, we investigate the track of automatic change detection. We propose a method lying on the registration and the over-segmentation of the images into superpixels. Then we describe each superpixel by its texture using texton histogram and its gray-level mean. A distance measure, such as Mahalanobis distance, allows to compare corresponding superpixels between two images acquired at different dates. We evaluate the performance of the proposed approach on images taken from the photographic landscape observatory produced during the construction of the French A89 highway.

Among the image segmentation methods we have tested for superpixel extraction, our experiments show the relatively good behavior of Achanta segmentation method.

The relevance of a change is strongly related to the intended application, we thus investigate a second track involving a user intervention. We propose an interactive change detection method based on a learning step. In order to detect changes between two images, the user designates with a selection tool some samples consisting of pixel sets

in “changed” and “unchanged” areas. Each corresponding pixel pair, i.e., located at the same coordinates in the two images, is described by a 16-dimensional feature vector mainly calculated from the dissimilarity image. The latter is computed by measuring, for each corresponding pixel pair, the dissimilarity of the gray-levels of the neighbors of the two pixels. Samples selected by the user are used as learning data to train a classifier.

Among the classification methods we have tried, experimental results indicate that random forests give the better results for the tested image series.

Keywords: image analysis, change detection, superpixels, texture, machine learning, classification.

Contents

List of Tables

List of Figures

Part I

Change Detection from Photographs

Chapter 1

Introduction

1.1 Objectives of the thesis

In this dissertation, we present our research work on change detection techniques and present new ways of classifying "changed" or "unchanged" areas for landscape images taken by digital cameras.

The main application domain of our studies is photographic landscape observatories (PLO): these databases, which are widely used in the field of integrated geography, consist of series of images obtained by strictly rephotographing the same landscape scene at regular time intervals.

The location of the tripod, the framing for composing the scene and the technical specifications of the device (sensor, lens, focal length, aperture) must be consistent between two images. This kind of setup has been formalized by French Ministry of Ecology, Energy, Sustainable Development and Land-use Planning in a 2008 document[1] . The landscapes are constantly changing, as a result of natural or human factors; it is necessary to follow these developments in order to understand it and to report on it: this is the main purpose of photographic landscape observatories. PLOs are used for short-term change analysis (impact studies for public development projects) as well as for long-term analysis (evolution of a landscape across the time); as such, they are decision-support tools. Last, France is part of The European Landscape Convention of the Council of Europe which promotes the protection, management and planning of European landscapes and organizes European co-operation on landscape issues[2] . PLOs are often seen as a way to contribute meeting the commitments of the Convention.

With a large number of documented spots and frequent updates, the number of images

[1] [Méthode de l'Observatoire photographique du paysage]
`http://www.developpement-durable.gouv.fr/IMG/DGALN_methodeOPP.pdf`

[2] [Florence Convention, 2000] `http://www.coe.int/en/web/landscape`

of a PLO quickly increases, and makes the human analysis very tedious. Furthermore, human analysis should be quite inaccurate when setting the boundaries of a region. So, a tool that automatically compares pairs of landscape photographs in order to highlight changes would be a great help for exploiting photographic landscape observatories.

Numerous studies have been conducted on change detection from satellite images (as we will describe in Chapter 2), which may be helpful to solve this problem. However, these existing methods cannot be directly applied to photographic landscape observatories. The main reasons come from 1) the utilization of classic digital cameras from the ground which raises some specific problems like the limitation of the spectral band number (i.e., low spectral resolution); 2) the strong variation of the depth in a same image which induces very different appearance of the same object categories depending on their position in the scene. In addition, two issues make change detection more difficult:

- the geometric misalignment of the images.
- the application-dependent significance of the changes: depending on the application, changes may be or not significant. For instance, the growth of vegetation on a naked soil is a significant change when studying vegetation recovery on a construction site, but it is non significant regarding seasonal changes in agriculture.

Fortunately, when collecting our image dataset, the setup (tripod location, framing of the scene and specifications of the camera) is quasi consistent between photos in the series. This property allows us to capture experimental images which are almost registered. But a geometric registration procedure is necessary to produce registered two-date images. However, lighting variations, seasonality, time of day induce completely different images at the pixel level, and thus lead to insignificant changes. Our goal is to design a system that would be robust to these insignificant changes and able to detect relevant changes of the scene.

1.2 Contributions and Outline

In this dissertation, our study mainly focuses on discriminating changed and unchanged regions in two images of the same scene taken by digital cameras at different times. We review the existing common approaches, and propose two new methods of change detection for photographic landscape observatories.

The dissertation is organized as follows.

Chapter 2 presents a review of the existing techniques. We mainly investigate image-algebra-based change detection techniques, transform-based change detection techniques, and classification-based change detection techniques.

In Chapter 3, we introduce an automatic change detection approach based on superpixel segmentation and texton encoding. This method considers not individual pixels but groups of pixels that represent meaningful homogeneous regions. The method contains three major processes: superpixel segmentation, texton encoding, and change detection. First, the preprocessed photographies are over-segmented in order to produce superpixels. Then each superpixel in the two-date image pair is represented using the texton histogram and gray-level mean. Finally, change detection analysis is performed by computing the dissimilarity between the two feature sets of the superpixel pair (one superpixel is in the first date image, and the other one is in the second date image). In our experiments, we compare two methods for superpixel over-segmentation using Mahalanobis distance for the comparison of the feature sets. Finally, the map of change level is thresholded into a binary map representing changed and unchanged regions.

The automatic change detection approach sometimes provides results which are unsatisfying for our application because some significant changes are not detected. One option is to ask for a user interaction to improve the detection results. Consequently, we propose in Chapter 4 an interactive change detection based on a dissimilarity measure and a supervised classification method. In this work, we introduce a discriminative model that is a 16-dimensional feature space comprising the textural appearance of dissimilarity image and features generated within a local neighborhood. Dissimilarity measure is used to detect the local differences in a neighborhood. To detect changes between two images, the user designates changed and unchanged samples in the images using a selection tool. These samples are used to train a classifier using a supervised method. The trained classifier is then applied to all the other pixels of the image pair in the testing step.

In Chapter 5, we conclude this dissertation and discuss some possible future directions.

Chapter 2

Change Detection Techniques

Abstract of the chapter

A change detection method that suits our needs must be able to identify "significant" changes while disregarding "non significant" changes. For a specific change detection project, one should design an algorithm that takes into account the different types of changes in order to make the distinction. MacLeod and Congalton [MC98] described the setup of change detection for monitoring natural resources: detecting if a change has occurred, identifying the nature of the change, measuring the area extent of the change, and assessing the spatial pattern of the change.

Because of the impact of multiple factors, it is not easy to determine which algorithm is suitable for a given application. Hence, a review of change detection techniques should be very useful to understand how these techniques work and can be use to solve various categories of problems. Consequently, this chapter reviews the existing change detection methods. These techniques mainly involve image preprocessing (e.g., image registration, radiometric adjustments), and change detection methods (e.g., image algebra, transformation-based change detection and classification-based change detection).

It is worth noting that since the techniques referring to remote sensing change detection address a similar problem to ours, we devote much of this review to them.

2.1 Preprocessing for Change Detection

Change detection consists in measuring or identifying pixel sets of significant changes by controlling non significant changes, such as differences caused by camera motion, sensor noise, illumination variation, non-uniform attenuation, and atmospheric absorption. To successfully perform change detection, it requires to carefully consider the environmental

characteristics and image processing methods. Some important environmental factors, such as atmospheric conditions, view angle, soil moisture conditions and phenological characteristics, should also be taken into account [Web01, Jen86].

Thus, before implementing change detection, a preprocessing procedure is necessary to satisfy the following requirements:

- Accurate registration of the two-date images.
- Accurate radiometric and/or atmospheric calibration.

In various preprocessing techniques, image registration and radiometric (e.g., intensity) adjustments are the most important ones. Then, we begin with the two preprocessing techniques followed by a review of change detection techniques.

2.1.1 Image Registration

Apparent intensity changes at a pixel caused by camera motion are never desired to be detected as real changes. Therefore, an accurate *image registration* is necessary for all change detection algorithms. Image registration consists in obtaining the alignment of several images into the same coordinate frame [RAAKR05]. If the objects of interest are almost rigid and the camera motion is small, we can use low-dimensional spatial transformation, such as affine, polynomial, elastic, similarity, or projective transformations to perform image registration. It is to note that registration error is one of the major factors resulting in change detection errors. Some researchers have studied the effects of registration errors, and they concluded that good registration performance is capable of leading to good change detection results [TJGM92, GLM92, DK98, Sto99, VB00, CFMC01]. Thus, the importance of accurate image registration for two-date image pair is obvious in change detection. Image registration techniques have been well studied and widely used in image analysis applications. Numerous good image registration methods have been constantly introduced [GSP86, Bro92, LMM95, ZF03, EM03, SSS08].

Eugenio and Marques [EM03] proposed a fully automatic geometric registration method for satellite images using a contour-matching technique. The method aims to extract the maximum reliable information of the image to perform the matching algorithm. The core premise of the method is the use of a contour-matching algorithm, which has three main steps: detection of the reliable areas and estimation of the gradient energy map; initialization of the contour positions; estimation of the transformation parameters of affine model. Thanks to the use of contour-matching model, this method has high accuracy for geometric correction of satellite images. Le Moigne [LMCC02] introduced an automatic parallel image registration for multi-spectral remote sensing images. The basic features of a correlation-based registration algorithm are generated by using the maxima of wavelet coefficients. Due to the use of high-frequency wavelet features, this method is similar to

an edge-based correlation method. Experiments proved that this method has higher accuracy and low complexity. Ding *et al.* [DTJ+10] proposed an image registration method using robust kernel principal component for object-based change detection. They used robust kernel principal component analysis method to capture the common pattern of the variform objects, then the variform objects can be precisely registered using robust kernel principal components.

2.1.2 Radiometric Adjustments

In general, the intensity variations caused by changes in the power or position of light sources in the scene are considered as "unimportant" changes. Even when there are no artificial light sources involved, there may be some physical effects with similar consequences on the image, such as calibration errors [RAAKR05]. Additionally, atmospheric conditions also influence the intensity variation. Relative radiometric normalization minimizes radiometric differences among images caused by inconsistencies of acquisition conditions rather than changes in surface reflectance [YL00]. Canty *et al.* [CNS04] used the multivariate alteration detection transformation to obtain invariant pixels for automatic relative radiometric normalization of multi-spectral satellite images. Since the apparent reflectance values of surfaces whose ground reflectance can be considered constant over time, the atmospheric effects can be estimated in a relative way by using the apparent reflectance values of surfaces. Based on this idea, Caselles and Lopez Garcia [CLG89] proposed a simple approach to atmospheric correction of multi-temporal data. Some studies used ground reference data or pseudo invariant features (PIFs) in the radiometric normalization of multi-temporal images [DTC02]. In this paper, the authors statistically selected PIFs, and used principal component analysis to find linear relationship between multi-temporal images. With this method satellite images can be normalized radiometrically and radiometric resolution for the multi-temporal images considered is unaltered. Hall *et al.* [HSNG91] have developed a technique to radiometrically correct multiple Landsat images of a scene to a reference image, by means of sets of scene landscape elements whose reflectance is nearly constant over time. The radiometric rectification algorithm provides a relative calibration when reliable atmospheric optical depth data or calibration coefficients are not available, otherwise, correction to absolute surface reflectance can be accomplished.

2.2 Review of Change Detection Techniques

Singh [Sin89] proposed that digital change detection approaches may be broadly characterized by the data transformation procedure (if any), and analysis techniques used to

delineate areas of significant alterations. Because digital change detection is affected by spatial, spectral, environmental, thematic and temporal constraints, the selection of a suitable method or algorithm for a given research project is important, but not an easy work.

To obtain high performance of change detection, the research of a reliable change detection technique thus becomes an important topic. During the past decades, a variety of change detection algorithms have been proposed [BS97, HHSMA98, BP00a, NC03, NC01, Wal04, IJ05, IM07, IJT08, BOC+13, COBG11, COB+14]. Previous reviews of change detection techniques have been collected by Singh [Sin89], Lu [LMBM04], Jensen [Jen05] and Radke [RAAKR05], which are all helpful surveys of change detection techniques. Some of these techniques include image algebra, multiple-date composite image change detection using principal component analysis, post-classification comparison, neural networks, expert systems, and spectral change vector analysis [Jen05]. Here, we refer to a relatively detailed categorization of change detection techniques presented by Lu *et al.* [LMBM04]. They grouped change detection methods into seven types: image algebra, transformation-based change detection, classification-based change detection, advanced models, Geographical Information System (GIS) approaches, visual analysis, and other approaches. In the following part, we will explore the first three of them that are very commonly used in practice. The key characteristics, the benefits and drawbacks, and some examples using these methods will be provided.

2.2.1 Image Algebra

The category of image algebra change detection includes image differencing, image ratioing, image regression, index differencing, change vector analysis (CVA) and background subtraction [LMBM04]. A common characteristic of such algorithms is the selection of thresholds to determine the changed areas. In this community, all the methods are relatively simple, easy to implement and interpret. But they cannot provide the complete matrices that represent the transition types of changes. One disadvantage of image algebra category is the difficulty in selecting suitable thresholds to identify the changed areas.

Macleod and Congalton [MC98] used Thematic Mapper (TM) data to make a quantitative comparison among three different change detection approaches including post-classification, image differencing, and principal component analysis. They found that the image differencing technique perform better than the post-classification and principal components analysis. The overall accuracy (i.e., the proportion of pixels correctly labeled) of change detection based on image differencing was 66%. Stow *et al.* [SCM90] found that the higher change detection accuracy can be obtained by ratioing multi-sensor,

multi-temporal satellite image data, which is a useful land use change enhancement technique, compared to principal components analysis. Sohl [Soh99] concluded that CVA can qualitatively result in rich details about the nature of a change by evaluating five methods including CVA, vegetation index differencing, univariate image differencing, a novel image differencing and post-classification differencing. Lyon *et al.* [LYLE98] compared seven vegetation indices for land cover change detection by evaluating the criterions: *1)* whether they lead to similar results; *2)* which statistical properties they have; *3)* whether they are feasible to detect the interesting changes. And they concluded that normalized difference vegetation index technique (NDVI) can provide the best performance of vegetation change detection.

1. Image Differencing

In the algebra change detection family, image differencing technique is the earliest to be used, and such approaches are still widespread. Early change detection methods were based on the signed difference image $D(x) = I_2(x) - I_1(x)$ between the first date image $I_1(x)$ and the second date image $I_2(x)$ at pixel x. The simplest model of image differencing can be denoted as

$$B(x) = \begin{cases} 1, & \text{if } |D(x)| > \tau \\ 0, & \text{otherwise,} \end{cases} \tag{2.1}$$

where $B(x)$ denotes the change map, and the threshold τ is chosen empirically. Although mathematically simple, image differencing allows for only one band information to be processed at a time [HS01]. And it is difficult to choose a suitable threshold. Many researchers gradually explore some new and advanced techniques for the selection of an interesting threshold. Rosin [Ros98] described four criterions that work on different principles for selecting thresholds. Smits and Annoni [SA00] discussed on how to choose a threshold in order to satisfy some requirements of false alarms and misses.

Another technique, similar to image differencing, is called image ratioing and defined as:

$$R(x) = \frac{I_1(x)}{I_2(x)}. \tag{2.2}$$

In the change detection results $R(x)$, unchanged pixels will have a value equal or very close to 1. Selecting an appropriate threshold is also a critical element in image ratioing change detection techniques.

2. Change Vector Analysis

In recent years, more and more studies demonstrate the advantages and potentials of change vector analysis. CVA is a conceptual extension of image differencing. This algorithm can detect all changes greater than a specified threshold, and can provide detailed change information. For the registered image pair, one can compute spectral change vectors, which describe the direction and magnitude of changes from date 1 image to date 2 image as shown in Figure 2.1, using spectral transformation and normalization. Transformations can improve the performance by adjusting for differences in atmospheric effect, and providing new spectral variables with dimensionality reduction, physical interpretability, and little or no loss of information. Normalization functions would gain importance as the scale or area of interest increased. Then, their magnitudes can be compared with a predefined threshold. Once that threshold is exceeded, the change is occurred. Moreover, one might use the other information contained in the change vectors, namely its direction, to know the types of changes [Mal80].

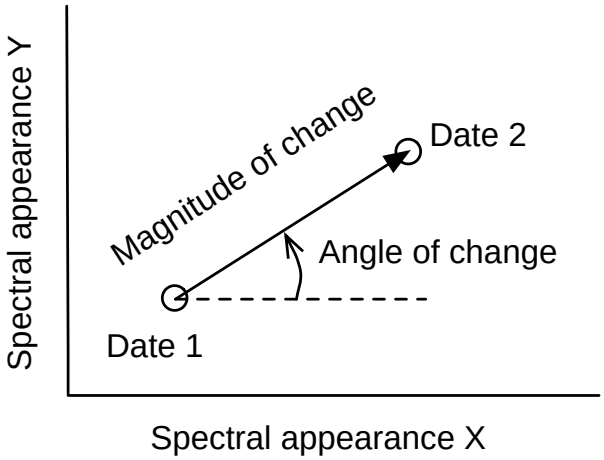


Figure 2.1: Spectral change vector.

Among the spectral change-based techniques, CVA has some advantages: not only it can prevent shortcomings of classification-based approaches such as cumulative errors in image classification for individual dates, but also it can use all bands to detect changes and provide "from-to" change information [CGH+03]. CVA is an effective land cover and land use change detection tool, which can be usefully involved in [JK10]:

- Process the full dimensionality of multi-spectral data in order to make the detection of all change present in the data.
- Extract the changed components in multi-spectral data.
- Facilitate the analysis of change images.

He *et al.* [HWS+11] proposed an extended CVA approach that incorporates textural change information into the traditional spectral-based CVA to detect land use and land cover changes. Bovolo and Bruzzone [BB07] proposed an unsupervised split-based approach (SBA) to make change detection in large-size multi-temporal images, which can handle both multi-spectral images and synthetic aperture radar (SAR) images.

2.2.2 Transformation-based Change Detection

The transformation-based change detection category includes principal component analysis (PCA), Tasselled Cap Transformation (K-T), Chi-Square, and Gramm-Schmidt (GS) transformations. Although these methods are able to reduce data redundancy between bands and highlight different information in components of interest, they fail to detailedly provide transition types of changes and require to select appropriate thresholds for detecting changes.

Compared to PCA transformation, K-T transformation coefficients are independent of the scenes. This advantage makes K-T useful in some change detection applications. Because of the relative complexity of Chi-Square and GS algorithms, they are less commonly used. In this section we mainly review the PCA change detection.

In the transformation category, PCA is most often used in detecting changes. Several studies have proved the ability of PCA technique in multi-temporal image analysis [MH94, CB96].

Theoretically, PCA consists of three steps [ES93]: *1)* calculate a covariance or correlation matrix using the input data sets; *2)* calculate the eigenvalues and eigenvectors; *3)* calculate the principal components. In change detection based on PCA, there are two ways to be considered:

- Independent data transformation, i.e., PCA is used to separately enhance multi-temporal images, and then, each image is separately classified and used in post-classification change detection.
- Merged data transformation, i.e., all the bands in the n-dimensional registered multi-temporal images are treated as a single N-dimensional data set input to PCA, where n is the number of bands in each image and d the number of dates, $N = n \times d$.

Moreover, the first two components tend to represent variations associated with unchanged information and overall image noise (e.g., atmospheric and seasonal variation), while the third and later components are of more interest in identifying changed information [BCM80]. Fung and LeDrew [FL87] successfully detect land-cover changes using minor components, and they indicated that the standardized PCA, i.e., the principal components are calculated using the correlation matrix, can improve the signal-to-noise ratio due to the ability of the principal components to identify interesting changes. Chavez

and Kwarteng [CJK89] presented a Selective Principal Component Analysis (SPCA), in which only two bands of the multi-date images are used as input data for PCA instead of all the bands. It can be used to detect and map spectral differences between different regions. With the selective PCA, they have successfully detected the changes of urban development and vegetation growth. Celik [Cel09] proposed a computationally simple yet effective, automatic and unsupervised change detection method using PCA and k-means clustering. The difference image is first segmented into $n \times n$ non-overlapping regions which are used to extract S ($S \leq n^2$) orthonormal eigenvectors by using PCA to construct an eigenvector space. Then, each pixel in the difference image is represented with a S-dimensional feature vector by projecting its $n \times n$ neighborhood data onto the generated eigenvector space. The feature vector space is clustered into two clusters by using k-means clustering. Finally, the change detection is achieved by assigning each pixel in the difference image to the one of the two clusters by using the minimum Euclidean distance between its feature vector and the mean feature vectors of the clusters.

2.2.3 Classification-based Change Detection

The classification-based change detection methods are based on the classified results, thus, the performance of classification is crucial for generating good change detection results.

1). Supervised and Unsupervised Change Detection

Supervised and unsupervised change detection algorithms have been proposed by [BS97, Sin89, HHSMA98, YB15, BP00a, VTCVK12, BOC$^+$13, VTB$^+$13]. The supervised classification methods require an available ground truth in order to derive a suitable training set for the learning process of the classifiers. However, the unsupervised methods performs change detection by making a direct comparison of the two multispectral images considered, without relying on any additional information.

***A.* Supervised Change Detection**

The supervised change detection technique classifies the images of the same scene by using supervised classification techniques. It aims at generating a change map, and recognizing the kinds of land cover transitions that have occurred. In the procedure of classification, an available ground truth and numerous training samples are required for the training step. The main advantage of supervised change detection is the robustness against the difference of atmosphere and environment, as well as the light conditions at the two acquisition times. We summarize this technique in Figure 2.2.

In this category, post-classification comparison is very commonly used. The traditional post-classification change detection firstly classifies the multi-temporal images into

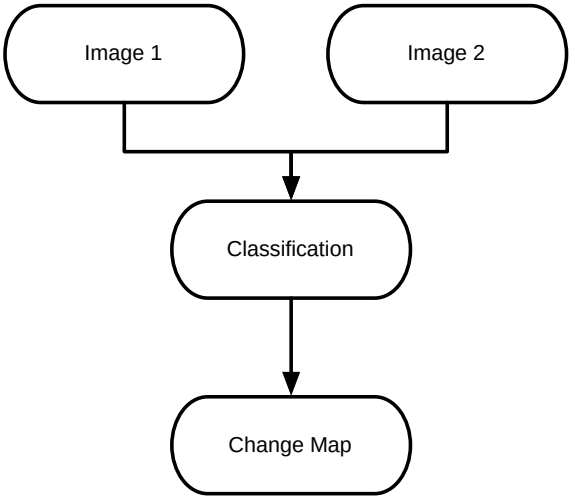


Figure 2.2: Main ideas of supervised change detection.

objects, then compares the classified images to extract the detailed "from-to" change information. Ahlqvist [Ahl08] developed a change detection method based on an extension of post-classification by using semantic similarity metrics. He used a fuzzy set-based approach to produce attribute-based prototype definitions of land cover classes to overcome class heterogeneity. The experiments demonstrated that his method can provide not only an overall and detailed land cover change information, but also an assessment of heterogenous land cover types of changes. Frate *et al.* [FPS08] introduced a change detection method to monitor and detect land cover changes from synthetic aperture radar (SAR) images based on post-classification logic. Since the use of single-channel SAR images makes difficult for achieving high classification accuracy, this method uses a multilayer perceptron neural network supervised classifier for classification of each SAR image, which provides the correct detection above 82%, missed alarm less than 18% and false alarms 0.3%.

In addition to supervised change detection based on post-classification techniques, many different supervised change detection approaches have been proposed. Based on the "compounding classification rule", Bruzzone *et al.* [BS97] proposed a supervised nonparametric technique to explicitly detect the land cover transitions that can be robust to different acquisition conditions at two different times. In addition, it is possible to implement change detection by using different sensors. This is a useful characteristic when change detection on a large temporal scale has to be performed and available images are provided by different sensors. Finally, by exploiting appropriate nonparametric classifiers, this method can use the multi-sensorial images acquired at two dates. Walter [Wal04]

introduced a change detection approach by using an object-based classification analysis of remote sensing data. This method classifies groups of pixels obtained by using a supervised maximum likelihood classification. In the classification, they construct a 16-dimensional feature set by calculating different measures of multi-spectral bands grouped by objects. The scheme improves the classification results, and thus helps to improve the performance of change detection. Generally, it is difficult to detect and classify the land cover transitions, especially for images with high within-class variance as well as low between-class variance. Therefore, Volpi *et al.* [VTB$^+$13] introduced an effective supervised change detection for very high geometrical resolution (VHR) images, by combining the benefits of the contextual information and the intrinsic properties of the Support Vector Machines.

The supervised change detection techniques are efficient, and exhibit some advantages over the unsupervised ones. They are:

- capable to identify the kinds of land cover transitions;
- robust to the changes of atmospheric and light conditions at the two acquisition times;
- capable to process multi-sensor or multi-source images.

Unfortunately, it is a difficult and expensive task to provide an appropriate ground truth in many applications, which is a practical limitation of the supervised methods. Consequently, when the ground truth is unavailable, the unsupervised classification for change detection will play an important role.

***B.* Unsupervised Change Detection**

The objective of unsupervised change detection is to produce a binary change map in which changed areas are separated from unchanged ones. The change detection is considered as an unsupervised classification problem, which consists of *1)* the comparison of the two-date images; *2)* the generation of "difference image"; *3)* the analysis of "difference image" to produce change map. The main ideas of unsupervised change detection techniques are summarized in Figure 2.3.

Hame *et al.* [HHSMA98] presented an unsupervised method to perform change detection and identification using clustering separately. First, clustering is performed on the two-date images to generate the "primary clusters". Then, the "secondary clusters" are produced within the primary clusters of the second date image. Finally, the change magnitude and change type can be identified by comparing the primary clusters in the first date image to the secondary clusters in the second date image.

Bruzzone and Prieto [BP00a] developed an adaptive parcel-based technique for unsupervised change detection. The method has the capability to adaptively extract the spatial-contextual information contained in the neighborhood of each pixel, which can efficiently make the effect of noise decreased and the change detection accuracy increased.

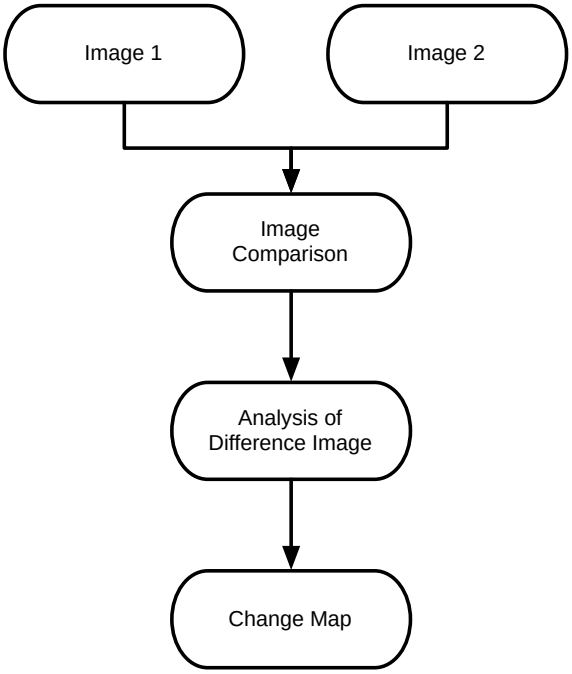


Figure 2.3: Main ideas of unsupervised change detection.

In particular, the definition of an adaptive pixel neighborhood can result in higher precision in the detection of boundaries of changed areas.

Boukir *et al.* [BOC+13] introduced a quasi-unsupervised (i.e., it is dependent on *test frames*, which only involves a few samples of small homogeneous areas in the image) region-based change detection in multi-spectral satellite images. An unsupervised classification in the high-resolution images benefitting from the post-classification scheme is used to detect subtle and important changes. As a result of a new spatio-temporal descriptor based on a relevant fragmentation rate, the method can differentiate "damaged" and "nodamaged" areas by using mean shift algorithm.

Yousif and Ban [YB15] proposed an unsupervised urban change detection. After a segmentation step, the method compares mean intensities of objects using an extension of the ratioing operation to generate a change image which is thresholded to produce a binary change map. It also proves the high effectiveness of the object-based technique in high-resolution SAR imagery.

Volpi *et al.* [VTCVK12] introduced an unsupervised kernel-based approach for change detection. This method based on a joint use of an initialization and an unsupervised cost function, performs change detection analysis by nonlinear clustering. The initialization routine provides a pseudo-training set, as well as the unsupervised cost function can optimize the parameters of the kernel.

In order to obtain both high change detection accuracy and efficiency, Castellana *et al.* [CDP07] present a new approach combining the advantages of supervised and unsupervised techniques. They used two different comparison strategies on the independently classified images to detect changes. If corresponding pixel pairs in the two images have sufficiently high posterior probabilities, the change analysis will be achieved by post-classification comparison, otherwise, a land cover transition matrix is automatically obtained from data by using an unsupervised change detection technique. Compared with traditional post-classification approach, the proposed method exhibits higher capability to detect real changes.

2). Expectation-maximization (EM) Change Detection

Expectation-maximization (EM) algorithm [DLR77] is an iterative method used in statistics for finding maximum likelihood or maximum a posteriori (MAP) estimates of parameters in statistical models, where the model depends on a set of unknown variables. The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate of the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. These estimations of parameters are then used to determine the distribution of the latent variables in the next E step.

In change detection research, the EM change detection is a classification-based method, which estimates the *a priori* joint class probabilities using EM algorithm in the two-date images. Some studies reported that EM change detection can give higher change detection accuracy than other methods. Bruzzone and Prieto [BP00b] developed two automatic techniques for the analysis of the difference image using unsupervised change detection. They proposed an iterative technique based on the EM algorithm, which allows the unsupervised estimation of the statistical distributions in order to discriminate the changed and unchanged pixels in the difference image. In the first technique, under the assumption that the pixels in the difference image are independent of one another, the selection of a decision threshold that minimizes the overall change detection error is automatic. In the second technique, in order to improve the accuracy of the final change detection results, it uses a Markov random field (MRF) algorithm that exploits the inter-pixel class dependency context to perform the analysis of the difference image.

In addition, Bruzzone and Prieto [BP02] presented a partially unsupervised classification approach in multi-temporal remote-sensing images. This approach investigates a novel definition of the EM algorithm according to the joint density function of image pairs. With this formulation, an unsupervised estimation can be performed between the prior joint probabilities of classes in the two-date images and the class-conditional density

functions in the second date image. In the process of estimation, the method has the ability of considering the additional prior information (if available) about the possible land cover changes, which allows a better estimation of detected changes.

3). Hybrid Change Detection

Hybrid change detection methods integrate the benefits of the threshold-based methods (e.g., image differencing) and classification-based methods. Such methods first identify the changed areas using the threshold-based methods, then classify and analyze the areas detected as changes. Although the methods can reduce classification error by excluding the unchanged pixels from classification, it involves the selection of a threshold for accurate classification.

Petit *et al.* [PSL01] presented a hybrid change detection method based on the combination of image differencing and post-classification. The method can detect detailed land-cover changes from produced change matrix. They found that such proposed approach can provide higher performance than post-classification techniques. Pixel-based classification methods seek to identify the class of each pixel in the image. Object-based schemes operate on groups of pixels, i.e., objects which are generated by using an appropriate image segmentation algorithm. Im *et al.* [IJT08] introduced a change detection method based on object correlation analysis and image segmentation. In this method, they selected three important features (called object correlation images, i.e., OCI) consisting of correlation, slope and intercept extracted from the objects by using an image segmentation technique. Compared to neighborhood correlation images (NCI) [IJ05], OCI were demonstrated to be more useful for detecting change in remote sensing data thanks to the benefits of object-based analysis. Aguirre-Gutiérrez *et al.* [AGSD12] presented a post-classification change detection method combining pixel-based and object-based land cover classification schemes. This approach can increase the accuracy both of the classification and change detection results. Niemeyer and Canty [NC03] investigated some pixel-based and object-oriented change detection for the routine nuclear verification using high-resolution imagery based on canonical correlation analysis (multivariate alteration detection MAD transformation). Their proposed method can enhance the change information in the difference images, and automatically select the significant thresholds using Bayesian techniques.

4). Artificial Neural Networks

Artificial neural networks (ANN) represent a nonparametric supervised method. ANN is able to estimate the properties of data based on training samples. The spectral data of the period of change can be used to train the neural network. Some studies have proved

that ANN is a meaningful tool for change detection [FPS08, WMPLC01, LLJ02]. It is worth noting that ANN is often sensitive to the size of the training set. In recent years, the studies of change detection based on ANN become more active.

Woodcock *et al.* [WMPLC01] demonstrated that ANN can be more effective in forest environments than traditional statistical approaches. In this research, they introduced a new approach that allows monitoring of large areas for forest change at frequent time intervals at reasonable expense due to the employment of generalization, which extends the geographic range and the temporal frequency of their applicability. Liu and Lathrop [LLJ02] presented an efficient approach based on ANN and PCA to detect newly urban areas using satellite images. Experimental results showed that the ANN-based method can lead to a 20%–30% more accurate detection than post-classification with requirement of only a modest training time. In addition, PCA can, not only reduce the computational cost, but also improve change detection accuracy in this method.

In addition to the methods reviewed above, some researchers introduced methods more computationally efficient. Gong [GSJC14] put forward a novel change detection technique in SAR images. This approach uses fuzzy C-means clustering to classify the changed and unchanged regions, and a novel MRF energy function is proposed to reduce the effect of speckle noise. In this method, all the steps involved are computationally simple.

Belghith *et al.* [BCA13] used kernel-based support vector data description to detect changes in multi-acquisition data. They proposed a new kernel function by incorporating some new information of the feature distribution and the dependencies among samples into the basic kernel functions. The copula theory which characterizes the dependencies among samples was first used in the support vector data description framework by them.

2.3 Conclusion

Due to the importance of monitoring changes of Earth's surface, research on change detection is a quite active topic, and a variety of change detection methods have been proposed. In this chapter, we first described the preprocessing for change detection, including the image registration and radiometric adjustments. The preprocessing for change detection is a very important step that helps reduce the influence of the difference of environmental and atmospheric conditions. Then we reviewed some widely used change detection methods of last decades, which mainly include three categories: image algebra change detection, transformation-based change detection, and classification-based change detection.

This overview presented theoretical background and different challenges inspiring our research. In the following chapters, we will propose two new change detection methods

partly inspired by some ideas presented in this chapter.

Chapter 3

Automatic Change Detection based on Superpixel Segmentation and Texton Encoding

Abstract of the chapter

In this chapter, we introduce an automatic change detection approach based on superpixel segmentation and texton encoding. In order to compensate the lack of information at pixel level, this method detects changes not from individual pixels but from groups of pixels, called superpixels, that represent meaningful homogeneous regions by means of an over-segmentation technique. The proposed method mainly consists of the following processes. Firstly, a homography-based image registration is used to produce registered image pairs. Then registered image pairs are over-segmented into superpixels. The main originality of our approach is the analysis of the characteristics of superpixels between two dates, in which we use texton histogram and gray-level mean to extract textural feature and intensity feature for each superpixel. Finally, change detection is performed by comparing the feature sets of corresponding superpixels in the two images. The dissimilarity between two feature sets of the two superpixels is measured by the Mahalanobis distance.

For extracting superpixels, we compare two segmentation algorithms: Shi segmentation and Achanta segmentation. For measuring the distance between the feature sets, we use Mahalanobis distance measure. The change matrix indicating the level of change of each superpixel is produced from the mean of the distance values between the feature sets of superpixel pairs. A binary change map is obtained by thresholding the change matrix with the use of a binarization technique.

Compared to the change detection based on the simple image differencing, our method

is able to improve the performance. Experiments show the relative efficiency and the capability of the proposed method for change detection.

3.1 Introduction

The application context of the work reported in this dissertation deals with the analysis of the temporal landscape evolution from image series taken from the ground, using a classical digital camera. Thanks to a tripod and a strict rephotography protocol in a same series (i.e., for a same scene), images are acquired in similar geometric conditions. So in order to preprocess the images, a homography-based image registration is sufficient to compensate for a possible small rotation of the camera.

Then, for a given registered image pair, the aim is to automatically locate change areas. Compared to most of the satellite images, our images have a low spectral resolution (at best three color channels). To take into account the lack of information at the pixel level, we propose to consider groups of pixels as the geometric primitives for change detection. In other words, we propose to partition the images into regions and to analyze their temporal evolution. The regions could be obtained by a classification of pixels into predefined semantic classes. But the use of a learning-based classifier hardly seems possible given the lack of already labeled images needed for training. Moreover, as we can see in the two examples given in Figure 3.1, a same object category, such as building class "BU" and tree class "TR", can have very different appearances in the image depending on the distance between the objects and the camera. Proportionally to the object size, this distance varies strongly in the scenes we considered. This noticeable difference from satellite images makes the use of classification or semantic segmentation challenging for our kind of landscape observatories.

It is therefore more realistic to move towards classical segmentation methods based on gray-level or color homogeneity. However, errors in the segmentation, like a region overlapping two objects, will induce errors in the change detection results. That is the reason why we investigate the use of over-segmentation methods giving some small regions called superpixels. One of the goals of these methods is that each superpixel overlays only a part of one object.

The principle of the proposed method lies on the analysis of the evolution of superpixel characteristics between the two dates. In order to be robust against insignificant changes caused by illumination, shadows or season, we choose to make use of the texture feature in addition to the gray-level mean of the images, since we think that the texture and intensity information can give a relevant description of each superpixel. Among various texture characterization methods, we chose texton encoding which has proved its potential. Then,

Figure 3.1: "'BU" denotes the building class, and "TR" denote the tree class.

two superpixels described by their texton histogram and gray-level mean can be compared by using a distance measure.

But all the pixels inside one superpixel in the first image may not belong to an individual superpixel in the second image. In other words, one superpixel in the first image can correspond to more than one superpixel in the second image. In order to estimate the level of change of one superpixel, more than one distance values have to be merged into a single value.

Finally, the binary change map can be obtained from the change matrix using a binarization technique.

The main steps of the proposed method are shown in Figure 3.2. Details are given in the following sections.

3.2 Geometric Image Registration

Even if the rephotography procotol is strict, the image pairs are not perfectly aligned. We proposed a homography-based geometric image registration procedure. Figure 3.3 summarizes the main steps of this preprocessing step.

We consider that the geometric relation between the two camera poses can be approximated by a rotation about the optical center. Under this assumption, the relation between the two images is a homography such that:

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad \lambda \in \mathbb{R}^* \tag{3.1}$$

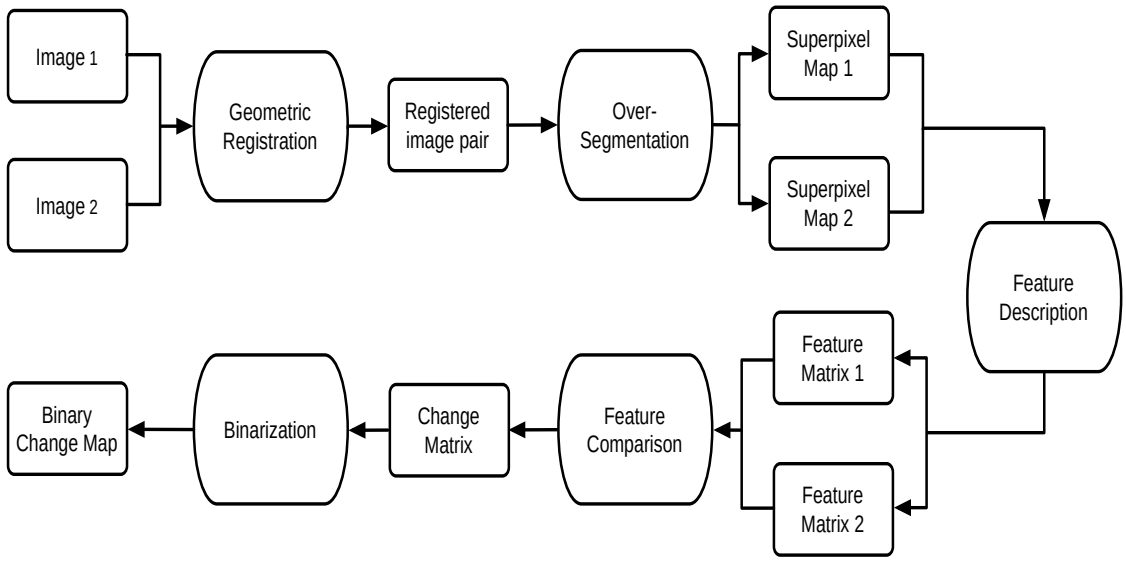


Figure 3.2: Flow chart of the automatic change detection method.

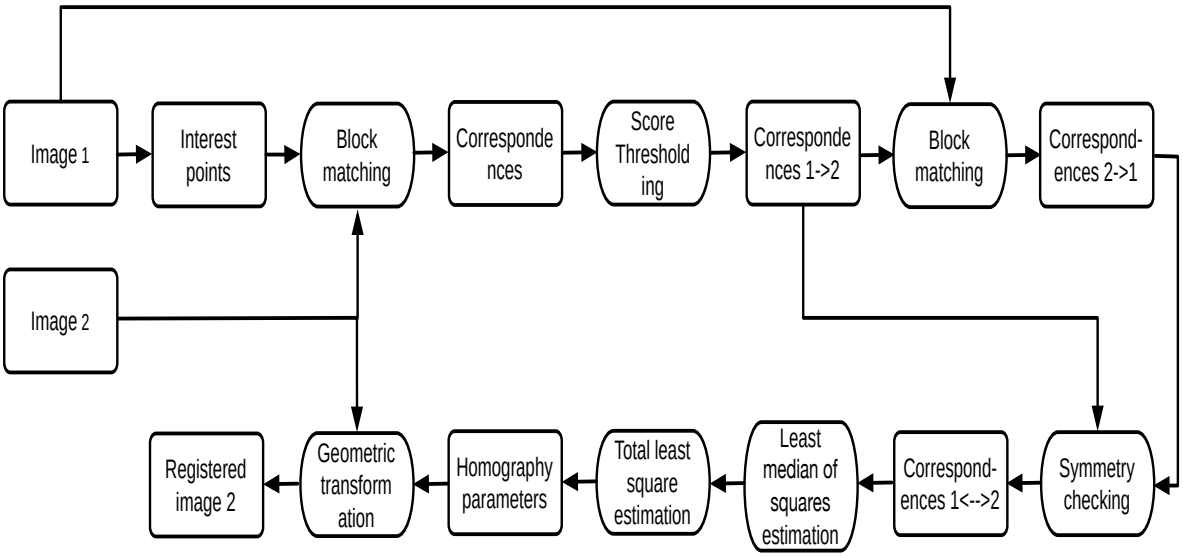


Figure 3.3: Geometric registration.

where (x, y) are the coordinates of a point in image 1, (x', y') are the coordinates of the corresponding point in image 2, and $\mathbf{H}$ is the 3×3 matrix containing the homography parameters.

In order to estimate the homography parameters, we can use point correspondences. In order to improve the alignment of the two images, these correspondence must be as accurate as possible. However, the large time interval between the acquisition of the two images induces a lot of changes making the establishment of reliable correspondences difficult.

Once reliable point correspondence are found, a classical parameter estimation technique can be used to calculate the homography parameters.

Finally, these parameters are used to transform one of the two images and to align them.

3.2.1 Finding Point Correspondences

Detection of Points of Interest

We extract points of interest in the first image using the method proposed by Shi and Tomasi [ST94]. For obtaining the response matrix of this detector, the two components $\mathbf{I}_x$ and $\mathbf{I}_y$ of the gray-level gradients are firstly estimated. We simply use the convolution with the 3×3 Sobel masks. Then, three matrices, $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{T}$, are computed from these gradient components. Elements located at row x and column y of these matrices are given by $\mathbf{P}(x, y) = {\mathbf{I}_x}^2(x, y)$, $\mathbf{Q}(x, y) = {\mathbf{I}_y}^2(x, y)$ and $\mathbf{T}(x, y) = \mathbf{I}_x(x, y)\mathbf{I}_y(x, y)$. Then, these three matrices are smoothed by a local mean filter. We use a 7×7 windows size.

The detector response matrix $\mathbf{R}$ is calculated from the 2×2 matrices $\mathbf{M}(x, y)$ defined for each pixel by:

$$\mathbf{M}(x, y) = \begin{bmatrix} \widetilde{\mathbf{P}}(x, y) & \widetilde{\mathbf{T}}(x, y) \\ \widetilde{\mathbf{T}}(x, y) & \widetilde{\mathbf{Q}}(x, y) \end{bmatrix} \tag{3.2}$$

where the tilde denotes the mean filtered versions of the matrices.

The response of a pixel (x, y) is the smallest eigenvalue of the symmetric matrix $\mathbf{M}(x, y)$:

$$\mathbf{R}(x, y) = \frac{1}{2}\left(\widetilde{\mathbf{P}}(x, y) + \widetilde{\mathbf{Q}}(x, y) - \sqrt{\left(\widetilde{\mathbf{P}}(x, y) - \widetilde{\mathbf{Q}}(x, y)\right)^2 + 4\left(\widetilde{\mathbf{T}}(x, y)\right)^2}\right). \tag{3.3}$$

Then, a non-maxima suppression is applied to the response matrix $\mathbf{R}$. We use a 7×7 neighborhood size.

Finally, the remaining non-zero responses are sorted such that we select the pixels corresponding to the strongest responses as points of interest.

Pixel Matching

Using a simple block matching algorithm, we find the corresponding pixels in the second image. We use zero-mean normalized cross correlation (ZNCC) measure:

$$\mathrm{ZNCC}(\mathbf{f}_1, \mathbf{f}_2) = \frac{(\mathbf{f}_1 - \overline{\mathbf{f}_1})^\top (\mathbf{f}_2 - \overline{\mathbf{f}_2})}{\|\mathbf{f}_1 - \overline{\mathbf{f}_1}\| \|\mathbf{f}_2 - \overline{\mathbf{f}_2}\|} \tag{3.4}$$

where $\mathbf{f}_1$ and $\mathbf{f}_2$ are the vectors of the gray-levels of the pixels in the two $N \times N$ windows whose similarity is measured, $\overline{\mathbf{f}_1}$ and $\overline{\mathbf{f}_2}$ are the means of the gray-levels.

We use a 7×7 window size and we reduce the search area to a 5×5 windows because the initial misalignment is small thanks to the strict rephotography protocol.

Selection of Reliable Correspondences

Some of the point correspondences are erroneous. So we keep only the most reliable correspondences by verifying two constraints: the matching score constraint and the weak bidirectional constraint.

The matching score (ZNCC value) measures the similarity between the gray-levels in the neighborhoods of the two pixels. The constraint requires that the score must be greater than a threshold. If the two neighborhoods are identical, ZNCC value is equal to 1. We use 0.75 for this threshold.

The weak bidirectional constraint states that when the order of the two images is inverted, the matching results must be identical. More precisely, if, during the correspondence search from image 1 to image 2, the pixel (x', y') is the correspondent of the pixel (x, y), then, during the correspondence search from image 2 to image 1, the pixel $(x \pm 1, y \pm 1)$ must be found corresponding to the pixel (x', y'). The ± 1 tolerance is used to take into account the discretization influence.

Sub-pixel Interpolation

The location of the remaining correspondences are refined to sub-pixel precision using the quadratic interpolation of the matching scores. If (x', y') is the corresponding pixel in image 2 and $s(x', y')$ denotes the matching score obtained for this pixel, then sub-pixel location (x_s, y_s) of the correspondent is given by:

$$\begin{cases} x_s = x' + \dfrac{s(x'-1, y') - s(x'+1, y')}{2\left[s(x'-1, y') - 2s(x', y') + s(x'+1, y')\right]} \\ y_s = y' + \dfrac{s(x', y'-1) - s(x', y'+1)}{2\left[s(x', y'-1) - 2s(x', y') + s(x', y'+1)\right]}. \end{cases} \tag{3.5}$$

Selection of Inliers

Despite the two constraints, some correspondences may be erroneous. We estimate the homography parameters using the least median of squares criterion for rejecting outliers [Rou84]. More precisely, given n corresponding pairs $(x_i, y_i) \leftrightarrow (x_i{}', y_i{}')$, $i = 1, \ldots, n$, we can write an over-constrained system of linear equations:

$$\mathbf{A}\mathbf{h} = \mathbf{0}, \tag{3.6}$$

with:

$$\mathbf{A}_{(2n\times 9)} = \begin{bmatrix} & & & & & \vdots & & & \\ x_i & y_i & 1 & 0 & 0 & 0 & (-x_i x_i') & (-x_i' y_i) & (-x_i') \\ 0 & 0 & 0 & x_i & y_i & 1 & (-x_i y_i') & (-y_i y_i') & (-y_i') \\ & & & & & \vdots & & & \end{bmatrix} \tag{3.7}$$

and

$$\mathbf{h} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{21} & h_{22} & h_{23} & h_{31} & h_{32} & h_{33} \end{bmatrix}^\top, \tag{3.8}$$

where $\mathbf{H} = \Big[h_{ij}\Big]_{\substack{i=1\cdots 3,\\ j=1\cdots 3.}}$

The least median of squares estimate of $\mathbf{h}$ is given by

$$\mathbf{h}_m = \underset{\mathbf{h}}{\operatorname{argmin}} \underset{i=1,\ldots,n}{\operatorname{med}} {r_i}^2 \tag{3.9}$$

where r_i is the i^{th} element of the residual vector $\mathbf{r} = \mathbf{A}\mathbf{h}$.

The estimation of $\mathbf{h}_m$ can be made using a Monte-Carlo type technique to draw q random samples of 4 correspondences.

The number q of samples is given by:

$$q = \frac{\log(1-P)}{\log\left(1-(1-\epsilon)^p\right)} \tag{3.10}$$

where p is the number of parameters ($p = 8$ for a homography), ϵ is the fraction of outliers and P is the probability that at least one of the q samples is outlier free. We use $\epsilon = 50\%$ and $P = 99\%$.

For each minimal sample of 4 correspondences, a vector $\mathbf{h}$ is calculated. The estimate of $\mathbf{h}_m$ is the vector $\mathbf{h}$ giving the smallest median m of the squared residuals. The residuals corresponding to $\mathbf{h}_m$ are then used to detect outliers. The i^{th} row of matrix $\mathbf{A}$ is removed if $|r_i/\widehat{\sigma}| > 2.5$, where $\widehat{\sigma} = 1.4826(1 + 5(2n - p))\sqrt{m}$. The term $1 + 5(2n - p)$ is used to compensate the effect of a small set of data. The constant 1.4826 compensates the bias induced by the median when data are also corrupted by Gaussian noise.

3.2.2 Estimation of the Homography Parameters

On one side, the least median of squares estimation is robust to outliers. On the other side, the least squares estimation is optimal in the presence of Gaussian noise. So, the homography parameters are finally estimated using the total least squares criterion from the inliers selected from the least median of squares estimation. The estimate of $\mathbf{h}$

$$\widehat{\mathbf{h}} = \underset{\mathbf{h}/\|\mathbf{h}\|=1}{\operatorname{argmin}} \|\mathbf{A}\mathbf{h}\|^2 \tag{3.11}$$

is the unit eigenvector associated to the smallest eigenvalue of the matrix $\mathbf{A}^\top\mathbf{A}$, where rows of $\mathbf{A}$ corresponding to outliers have been removed.

3.2.3 Image Registration

The homography parameters are finally used to transform image 2 using bilinear interpolation. Last, the final registered image pair is obtained by calculating the maximum-area axis-aligned rectangle included in the intersection of image 1 and transformed image 2 as illustrated in Figure 3.4.

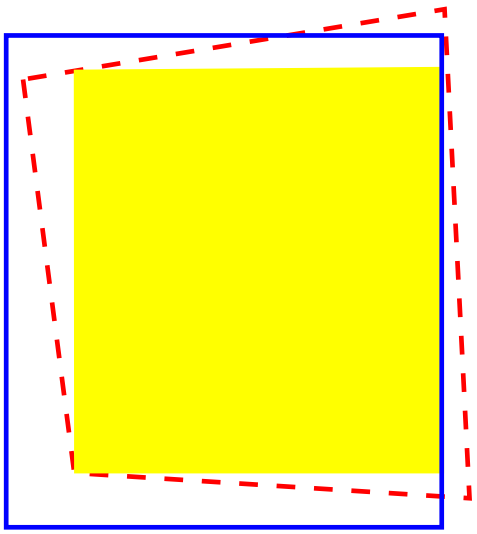


Figure 3.4: Common area of a registrated image pair: blue rectangle represents image 1, red dashed quadrilateral is the transformed image 2 and yellow filled rectangle is the maximum-area axis-aligned rectangle included in the intersection of image 1 and transformed image 2.

3.2.4 Example

An example of such transformed image pair is given in Figure 3.5. To obtain this result, we requested 1000 points of interest. After applying the matching score constraint, 156 correspondences remain. The weak bidirectional constraint left 119 correspondences. Finally, the outlier rejection using least median of squares estimation gave 96 correspondences shown in Figure 3.6. The initial registration error given by the mean Euclidean distance between the 96 correspondent points was equal to 3.35 pixels. After the final parameter estimation and the registration, this error is reduced to 0.88 pixel.

Figure 3.5: Example of a registered image pair.

3.3 Superpixel Extraction using Over-Segmentation

As we described in the introduction of this chapter, in order to reduce errors in the segmentation results consisting of regions overlapping more than one object, we use an over-segmentation method to extract superpixels, and our goal is to try to make sure that each superpixel overlays only a part of one object.

In this study, we first investigate two state-of-the-art superpixel segmentation algorithms, including an efficient graph-based image segmentation using normalized cuts (i.e., Shi segmentation) [SM00], and Simple Linear Iterative Clustering (SLIC) superpixels (i.e., Achanta segmentation) [ASS+12, ASS+10]. These segmentation algorithms will be later experimented in our proposed approach.

Figure 3.6: Examples of point correspondences: 96 correspondences used to obtain the registered image pair shown in Figure 3.5.

3.3.1 Shi Segmentation

Shi segmentation [SM00] is based on a graph representation of the image where each pixel is a node. The weight of an edge between two nodes reflects the likelihood that the two pixels belong to one object. This weight depends on the spatial locations and the gray-levels of the two pixels. The algorithm relies on a recursive bipartition of the graph minimizing a new criterion called normalized cuts which measures the total dissimilarity between the regions and the total similarity within the regions. The minimization of this criterion comes down to solve a generalized eigenvalue problem.

3.3.2 Achanta Segmentation

Achanta segmentation [ASS+12, ASS+10] generates superpixels by clustering pixels using their color similarity and proximity in the five-dimensional space $[l\ a\ b\ x\ y]$. Here $[l\ a\ b]$ is the color vector of the pixel in CIELAB color space, which is widely considered as perceptually uniform for small color distances, and $[x\ y]$ is the location of the pixel. This is done by introducing a new distance measure D_{labxy} that is defined as follows:

$$D_{labxy} = D_{lab} + \frac{m}{S} D_{xy} \tag{3.12}$$

where

$$D_{lab} = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2},$$
$$D_{xy} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}, \tag{3.13}$$

and $S = \sqrt{N/K}$, where N is the number of pixels and K the desired number of approximately equally-sized superpixels.

The distance measure D_{labxy} represents the sum of the *lab* distance D_{lab} and the *xy* plane distance D_{xy}. The constant m introduced in D_{labxy} allows to control the relative importance between color similarity and spatial proximity.

3.4 Superpixel Description

In addition to the gray-level mean, we propose to describe each superpixel by its textural appearance to be more robust against "non significant" changes caused by shadows and seasons.

Texton, introduced by Julesz [Jul81], is analogous to a phoneme in speech recognition. He qualitatively described it for simple binary line segment stimuli: *oriented segments, crossings and terminators*, but did not provide an operational definition for gray-level images. Subsequently, Malik *et al.* [MBLS01] re-invented textons as frequently co-occurring combinations of oriented linear filter outputs. Algorithmically, each texture is analyzed using a filter bank, and each pixel is transformed to an N-dimensional vector of filter responses, where N is the real-valued filter responses. These vectors can be clustered using k-means approach. The criterion is to find k centers such that the sum of the squared distance from the centers are minimized after assigning each data vector to the nearest center. k-means is a greedy algorithm which iteratively performs the following two operations:

- Assign data vectors to the nearest of the k centers.
- Update each of the k centers to the mean of the data vectors assigned to it.

These two steps are iterated until the algorithm converges and a local minimum of the criterion is achieved. These centers are the *textons*.

By assigning each pixel to its nearest texton, each superpixel can be described by the normalized histogram of textons. In our experiments, we use $k = 32$ textons. Adding its gray-level mean, each superpixel is thus described by a 33-dimensional feature vector.

3.5 Distance between the Feature Vectors

Distance measures play an important role in image processing. Throughout the years, many such measures have been proposed for measuring the dissimilarity between the feature sets, but it is very important to make a choice that depends on the application considered. The most commonly used measures contain Euclidean distance, Chi-square distance, Hamming distance, and so on.

To describe each superpixel, the feature vector we use consists of the texton part, i.e., normalized histogram values, and the gray-level mean. These components are of a different nature and, in order to avoid a greater influence of a component with a greater scale when comparing two feature vectors, there is a need of a kind of normalization. Moreover, the dependence between components deserves to be taken into account.

With the underlying assumption of a multivariate Gaussian distribution, Mahalanobis distance is able to take int account different variances in each component and the covariance between components. It can be seen as the transformation of the data into a unit variance uncorrelated data before computing the Euclidean distance.

For two feature vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, where $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)$, and $n = k + 1$ is the number of features, Mahalanobis distance is defined by:

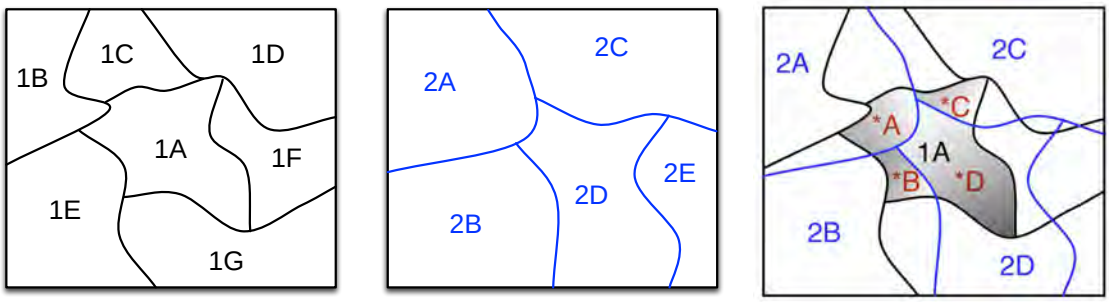
$$d_m(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(\boldsymbol{x} - \boldsymbol{y})^{\mathrm{T}} \boldsymbol{S}^{-1} (\boldsymbol{x} - \boldsymbol{y})}, \tag{3.14}$$

where $\boldsymbol{S}^{-1}$ is the inverse of covariance matrix $\boldsymbol{S}$.

3.6 Superpixel Evolution Assessment

An important step in the comparison of the feature sets is how to extract the superpixel pair. Assume we have two images: the first date image I1 and the second date image I2. As shown in Figure 3.7, the boundaries of all superpixels for I1 and I2 are plotted as black and blue lines, respectively. Because of the significant and non significant changes, the segmentation results of the two images are different. So, one superpixel in I1 might overlap parts of several superpixels in I2. For example, the superpixel "$1A$" in I1 overlaps the superpixels "$2A$", "$2B$", "$2C$" and "$2D$" in I2. Then, we denote "$*A$" as the intersection of one superpixel pair i.e., $*A = (1A, 2A)$. Similarly, "$*B$", "$*C$", and "$*D$" are all the superpixel pairs corresponding to superpixel "$1A$" in I1. Thus, the superpixel pairs corresponding to superpixel "$1A$" can be described by $\{*A, *B, *C, *D\} = \{(1A, 2A), (1A, 2B), (1A, 2C), (1A, 2D)\}$.

As described above, one superpixel in I1 may correspond to p, $p \geq 1$, superpixels in I2. For each of these p correspondences, Mahalanobis distance between the feature vectors



(a) (b) (c)

Figure 3.7: Superpixel evolution assessment: superpixels extracted from (a) date 1 image and (b) date 2 image; (c) superpixel correspondences associated to superpixel 1A.

is calculated. We estimate the level of change of the superpixel in I1 by calculating the mean value among the p Mahalanobis distances.

In the exemple shown in Figure 3.7, the level of change of superpixel "1A" is given by:

$$ch(1A) = \text{mean}\{d_m(1A, 2A), d_m(1A, 2B), d_m(1A, 2C), d_m(1A, 2D)\} \tag{3.15}$$

where $d_m(S_1, S_2)$ denotes the Mahalanobis distance between the feature vectors describing the two superpixels S_1 and S_2.

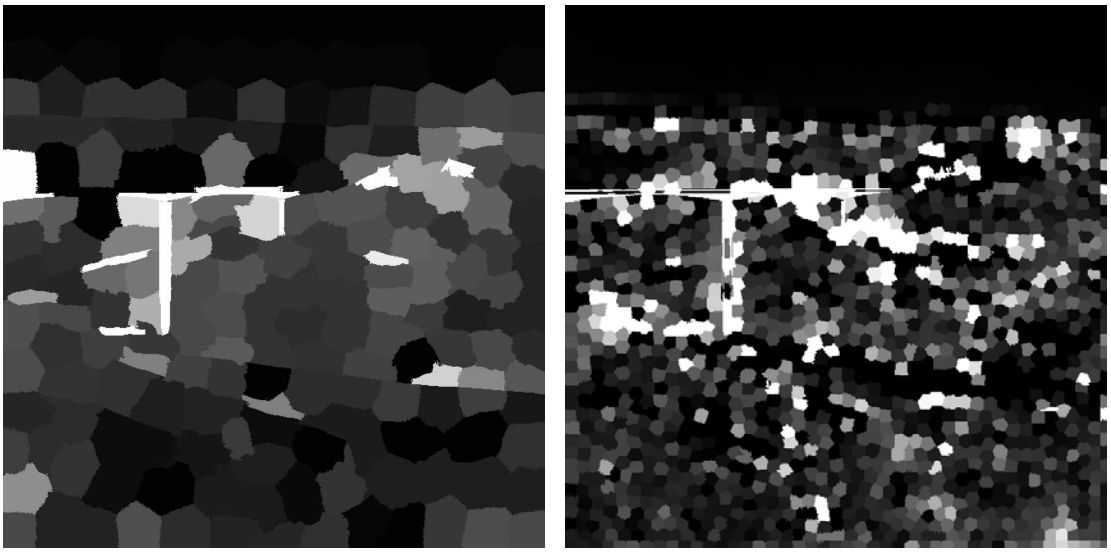
The covariance matrix is previously computed from the feature vectors describing all the superpixels in the image pair.

Then, each element in the change matrix corresponding to pixels belonging to the superpixel in image I1 is assigned its change level. The similar process is applied to the other superpixels in Image I1. Once all the processes are finished, the change matrix is obtained.

3.7 Computation of the Binary Change Map

As described in Section 3.6, for an image pair, the change level of each superpixel in the first image is assigned to the corresponding pixels, giving a change matrix. This matrix can be converted by an affine transformation into a gray-level image. Figure 3.8 shows two images representing two change matrices obtained using two over-segmentation methods and the image pair presented in Figure 3.5.

The last step of our method is to decide whether each pixel has changed or not from its change level. We propose to binarize the gray-level image obtained from the change matrix using a global threshold. In the experiments described below, we use a threshold that optimizes the F1 score of the binary change map. This criterion, detailed in



(a) (b)

Figure 3.8: Change matrix obtained using (a) Shi segmentation and (b) Achanta segmentation.

Section 3.8.2, can represent the global performance of the change detection method.

3.8 Experiments and Discussions

3.8.1 Study Site and Database

The images used in this study come from a collection of photographs taken during the construction of the French highway A89 as shown in Figure 3.9.

The whole project has been documented before, during and after the construction process, and this comprehensive database forms a photographic landscape observatory ([EG06]). We worked on 19 images taken at 7 locations. The size of each image is about 650×650. In every location, the specific series of images was taken once a year. Figures 3.10 and 3.11 display some images acquired between January, 1999 and March, 2006.

Our proposed automatic change detection method was evaluated using all of the 19 images. In order to obtain better visual comparison, we sort all images into 12 image pairs, and each image pair consists of one image before the construction of the highway and the other one during the construction.

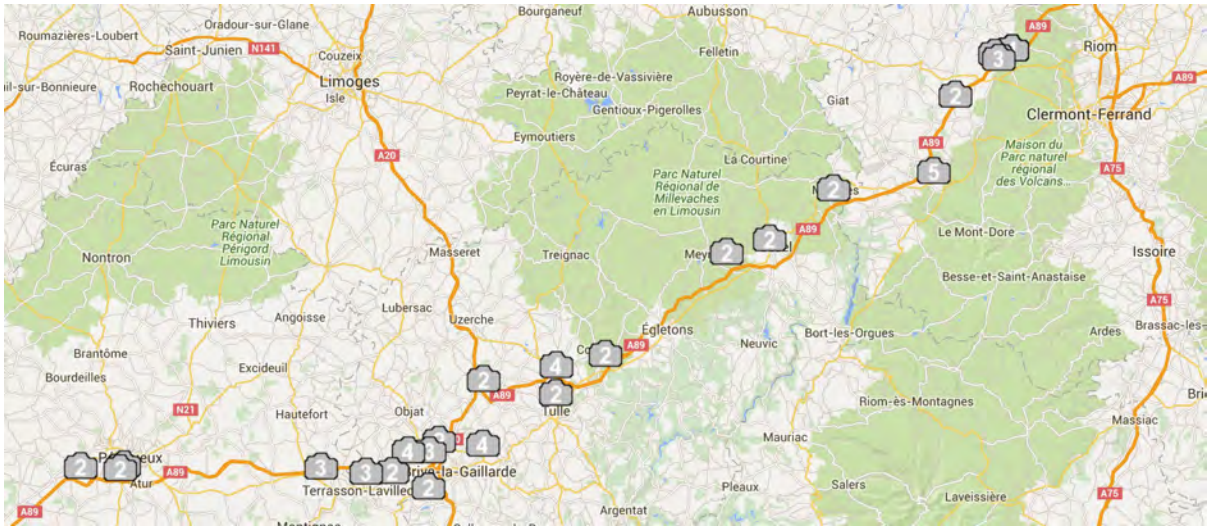


Figure 3.9: Photographic landscape observatory of the highway A89.

3.8.2 Performance Evaluation

The performance of a change detection method can be evaluated visually and quantitatively based on the application needs. For the quantitative assessments, we use hand-generated reference data (ground truth) obtained with an image editor. In our project, ground truth images involve the assigned colors acting as indices of the two classes, where white denotes changed and black denotes unchanged, respectively. Note that due to visual ambiguity, it is quite difficult to generate perfect ground truth change maps. The usual difficulty to delineate accurately the objects is increased by the presence of strong illumination changes and shadows.

The simplest accuracy measure is overall accuracy, i.e.,the proportion of pixels correctly labeled. We are also interested in two other measures: omission error and commission error. Omission error occurs when pixels belong to a class but fail to be labeled as such. Omission error produced in changed regions ($O_{changed}$) and in unchanged regions ($O_{unchanged}$) can be defined as:

$$O_{changed} = \frac{FN}{TP + FN} \qquad O_{unchanged} = \frac{FP}{FP + TN} \tag{3.16}$$

where TP is the number of pixels correctly labeled as changed (true positives); FP is the number of pixels wrongly labeled as changed (false positives); FN is the number of pixels wrongly labeled as unchanged (false negatives); TN is the number of pixels correctly labeled as unchanged (true negatives).

Commission error occurs when pixels that belong to the other class are wrongly labeled as belonging to the class. Commission error produced in changed regions ($CO_{changed}$) and

Figure 3.10: Examples of image series (part 1).

Figure 3.11: Examples of image series (part 2).

in unchanged regions ($CO_{unchanged}$) can be defined as:

$$CO_{changed} = \frac{FP}{TP + FP} \qquad CO_{unchanged} = \frac{FN}{FN + TN}. \tag{3.17}$$

Another commonly used measure is F1 score, which considers both the precision and the recall of the classification results. Precision (denoted by p) is the positive predictive value, i.e., the number of pixels correctly labeled as changed divided by the number of pixels labeled as changed. Recall (denoted by r) is the true positive rate, i.e., the number of pixels correctly labeled as changed divided by the number of changed pixels. F1 score is the harmonic mean of precision and recall, and is defined by:

$$F_1 = \frac{2pr}{p + r} \tag{3.18}$$

where $p = \frac{TP}{TP + FP}$ and $r = \frac{TP}{TP + FN}$.

F1 score is a value between 0 and 100% (100% when the classification result is identical to ground truth).

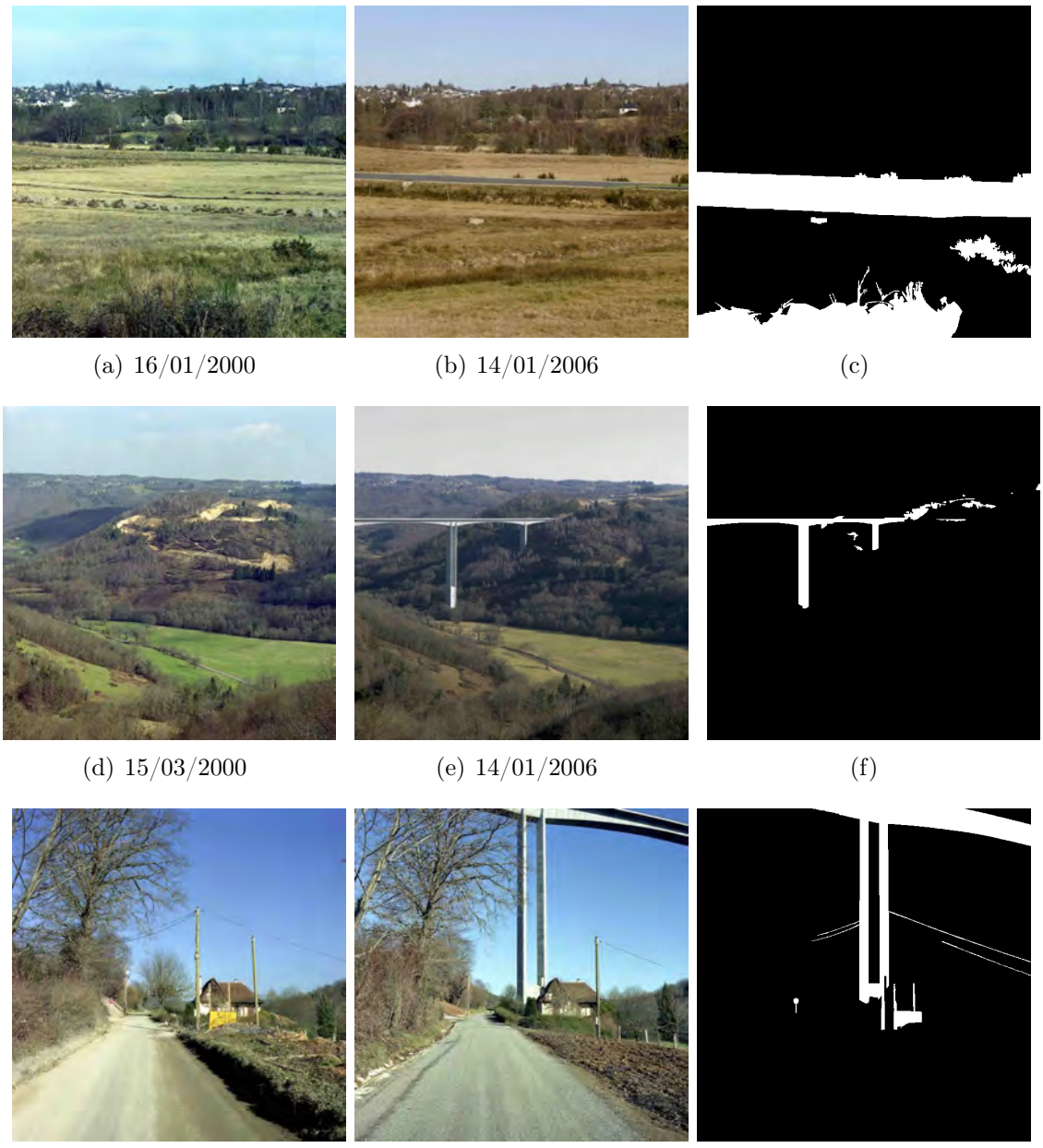
3.8.3 Experimental Results using Shi Segmentation

In this subsection, we present the experimental results on three image pairs shown in Figure 3.12. The segmentation results obtained using Shi method are given in Figure 3.13. Finally, the change detection results based on this segmentation are presented in Figure 3.14.

We perform the Shi segmentation using the open source code provided by the author[1] . It can be seen in Figure 3.13 that Shi segmentation generates regular superpixels, but sometimes it does not adhere to object boundaries. As shown in Figure 3.13(d), it can not produce superpixels exactly along the boundary of an object. For example, the produced superpixels near the boundary of the bridge always contain more than one object (e.g., superpixel "P"). This error directly impacts the change detection results, thus, leads to incomplete change detection as seen in Figure 3.14(d). That is, part of region "P" is detected as unchanged in the binary change map, while the region "P" belonging to the bridge should be a changed region.

We can see for example the effects of segmentation errors inducing incomplete detection of changed pixels (e.g., "P", "Q1" and "Q2" on the bridge and "Q3" on the highway) or regions wrongly detected as changed (e.g., "Q4" on the tree branches with the sky in the background).

[1] The code is available at `http://www.cis.upenn.edu/~jshi/software/`

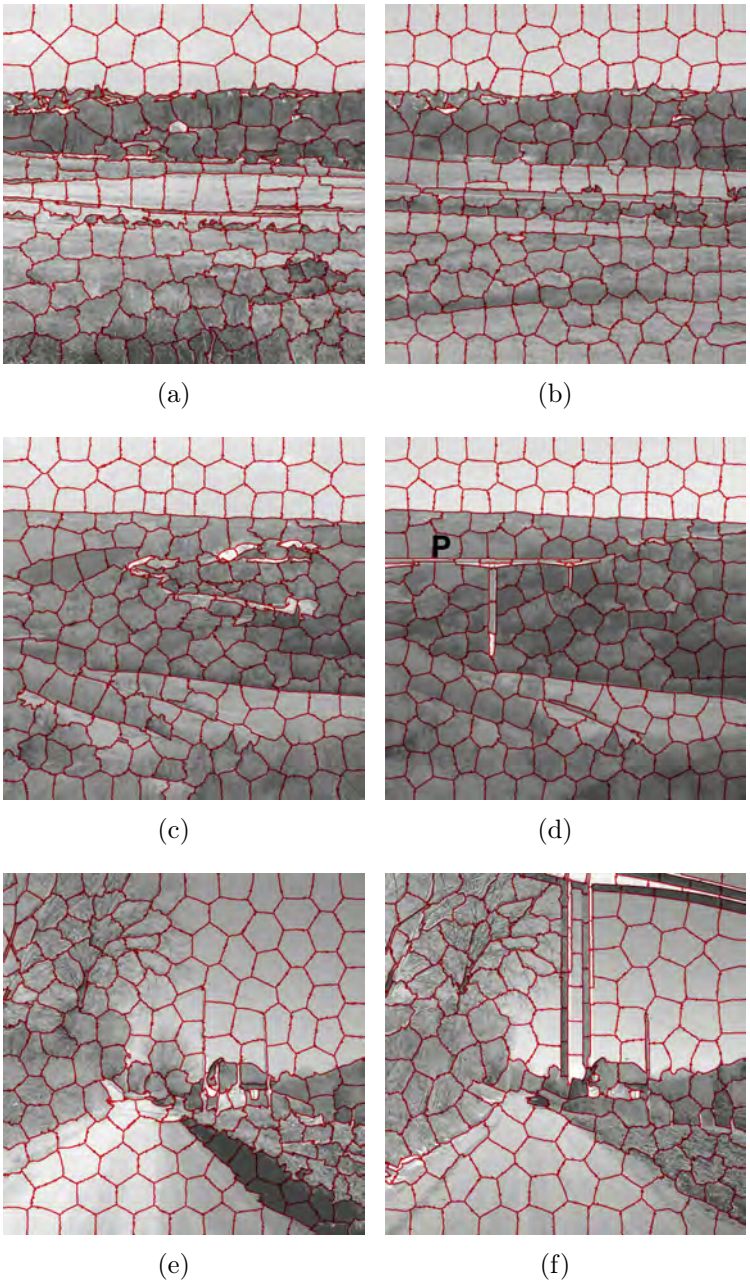


(a) 16/01/2000 (b) 14/01/2006 (c)

(d) 15/03/2000 (e) 14/01/2006 (f)

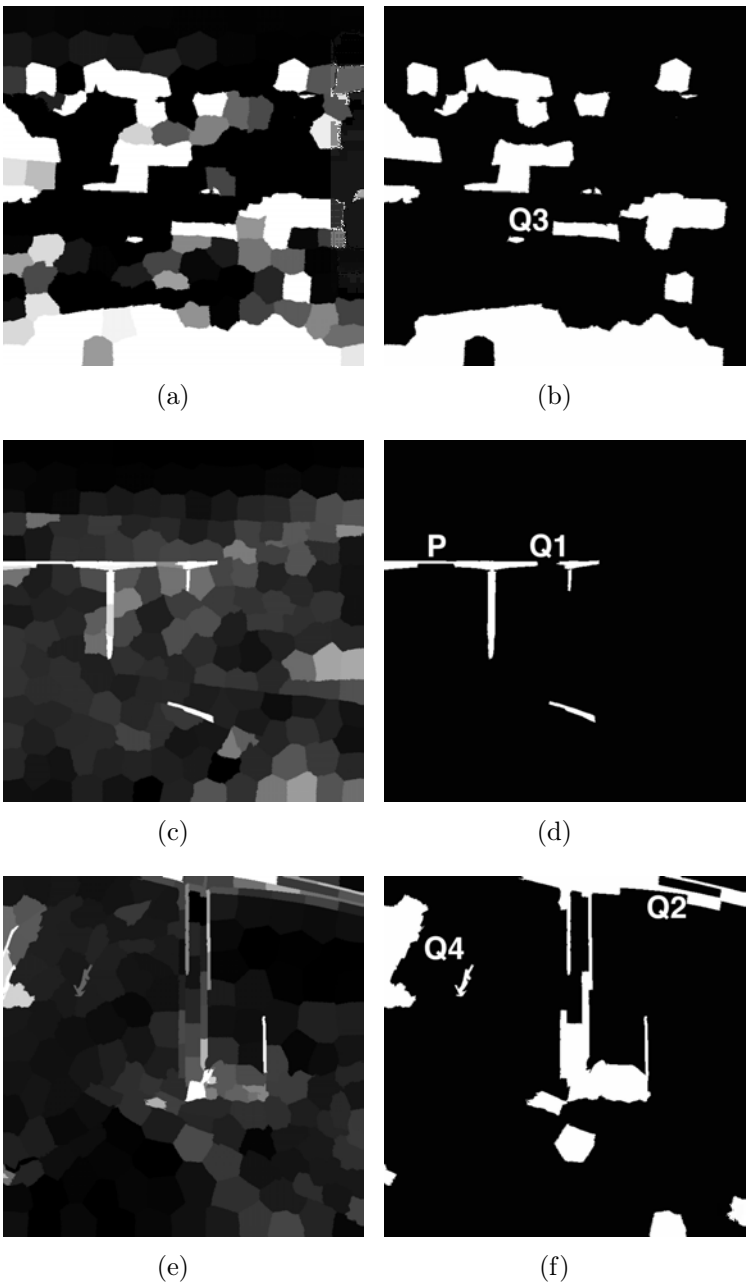
(g) 15/02/2001 (h) 24/01/2003 (i)

Figure 3.12: Examples of image pairs with ground truth: (a), (d), (g) first and (b), (e), (h) second date images; (c), (f), (i) ground truth change maps.



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

Figure 3.13: Shi segmentation results of image pairs presented in Figure 3.12.



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

Figure 3.14: Change detection results with Shi segmentation: (a), (c), (e) change matrices; (b), (d), (f) final change maps.

Table 3.1 presents performance evaluation, which is obtained by calculating the mean, standard deviation and median values of different performance measures considering the 12 image pairs of our data set. Although the mean overall accuracy is higher than 80%, the mean F1 score is only 40.35%. Observing the change detection results, we can find that some subtle changes can not be extracted, and some areas are wrongly detected. Thus, as to capture the changed pixels, this method using Shi segmentation gives high mean values of omission error with 53.67% and commission error with 58.76%. For distinguishing unchanged pixels, mean omission error is higher than 16%; lower mean commission error with 8.62% indicates that less pixels are wrongly detected as unchanged. The high standard deviation values in Table 3.1, show the high dispersion among different values in the same performance measure. We thus conclude that this method using Shi segmentation does not provide satisfying change detection results for all the image pairs, although the median overall accuracy is high with 93.69%.

The change detection using Shi segmentation gives relatively poor performance. In addition, it is worth noting that the runtime is high. The size of our experimental images is about 650×650, the runtime lasts about $700s$ per image using a cmex Matlab implementation on a 2.5 GHz Core i5.

Table 3.1: Performance evaluation of the automatic method with Shi segmentation.

	Overall accuracy	Omission error		Commission error		F1 score
		changed	unchanged	changed	unchanged	
Mean [std[1]]	83.74% [21.00%]	53.67% [16.45%]	16.82% [26.06%]	58.76% [20.06%]	8.62% [13.46%]	40.35% [15.77%]
Median	93.69%	52.83%	4.39%	59.74%	2.07%	45.02%

[1] "std" denotes standard deviation.

3.8.4 Experimental Results using Achanta Segmentation

In this subsection, the change detection experiments using our proposed automatic method based on Achanta segmentation are described. We use the same image pairs as shown in Figure 3.12(a)(b), (d)(e), and (g)(h). The ground truth images are shown in Figure 3.12(c)(f)(i).

We perform the Achanta segmentation algorithm using available source code embedded in the state-of-the-art software VLFeat[2] . In our tests, we set to 15 the parameter giving the starting size of the superpixels. The other parameter "*regularizer*" is set to 0.3,

[2] The code is provided in `http://www.vlfeat.org/download.html`

which is the trade-off appearance for spatial regularity when clustering (a larger value results in more spatial regularization). The segmentation results of the image pairs are presented in Figure 3.15. Among the two over-segmentation methods considered, Achanta segmentation outperforms Shi segmentation. It is the fastest method, over-segmenting our experimental images in less than 2s, and it is memory efficient.

Figure 3.16 shows the change detection results based on Achanta segmentation, where Figure 3.16 (a), (c) and (e) are the change matrices for each image pair, and Figure 3.16 (b), (d) and (f) are their binary change maps. From the change detection results, we can see not only the desired changed areas, like the bridges, are successfully extracted, but also some subtle changed regions can be also correctly detected, such as "T1" and "T2" in Figure 3.16 (d) and (f). These results show a satisfactory behavior in the presence of shadows, and the runtime is relatively short. However, some small regions like the leaves "T3" are wrongly detected as changed.

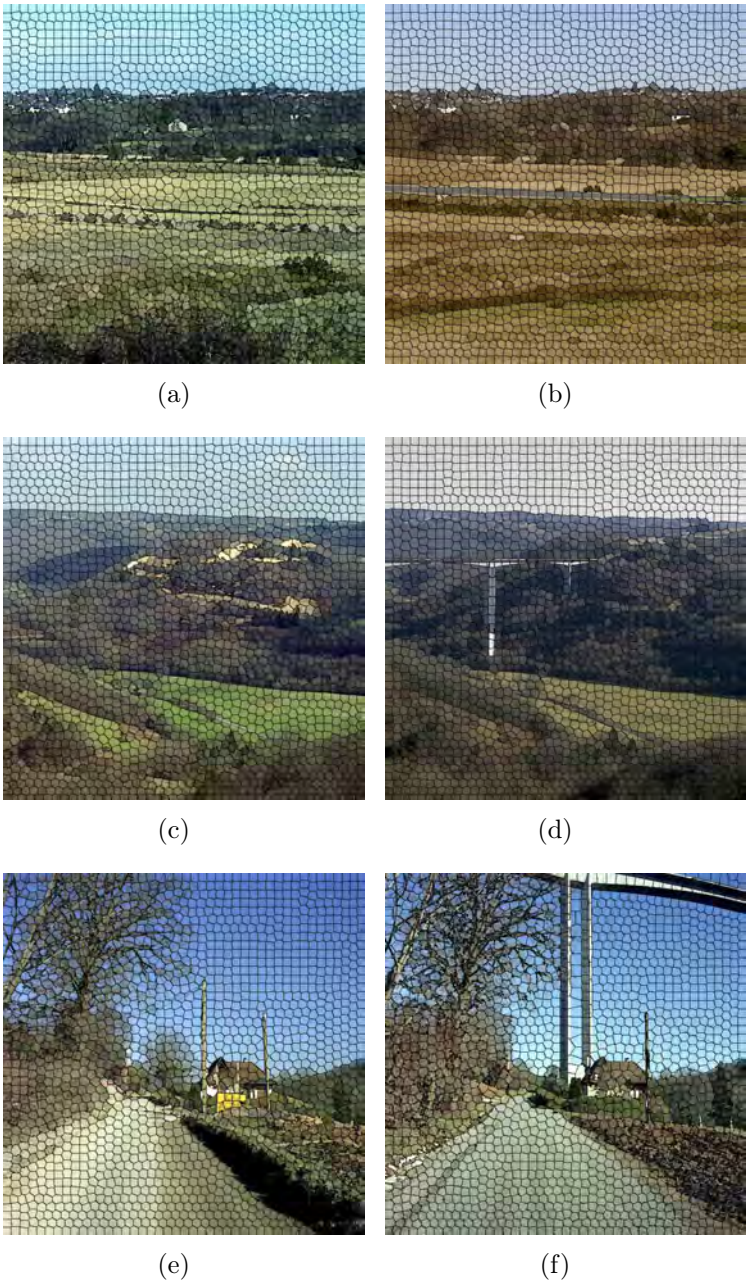
Table 3.2 evaluating the performance of the method is produced by calculating the mean, standard deviation and median values of different measures for all the image pairs. The mean overall accuracy is higher than 82%, and the mean F1 score is 41.40%, which are better than the experimental results using Shi segmentation. For detecting the changed pixels, the mean omission error is 43.89% which is decreased by 9.78% than using the Shi segmentation. Unfortunately, the mean commission error is still high with 62.54%. Although the standard deviation values of different measures are relatively decreased, the change detection method using Achanta segmentation cannot generate good results for all image pairs, thus, it is a challenging work for us to improve the automatic change detection method.

Table 3.2: Performance evaluation of the automatic method with Achanta segmentation.

	Overall accuracy	Omission error		Commission error		F1 score
		changed	unchanged	changed	unchanged	
Mean [std]	82.83% [21.57%]	43.89% [22.61%]	16.08% [24.26%]	62.54% [19.19%]	5.63% [10.39%]	41.40% [15.17%]
Median	95.50%	46.18%	3.25%	63.57%	1.70%	44.70%

3.8.5 Comparative Experiment

In order to better evaluate our method, we compare a change detection method based on the simple image differencing [Sin89] with ours. In this method, firstly a dissimilarity image is calculated using image differencing, then, the dissimilarity image is thresholded



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

Figure 3.15: Achanta segmentation results of image pairs presented in Figure 3.12.

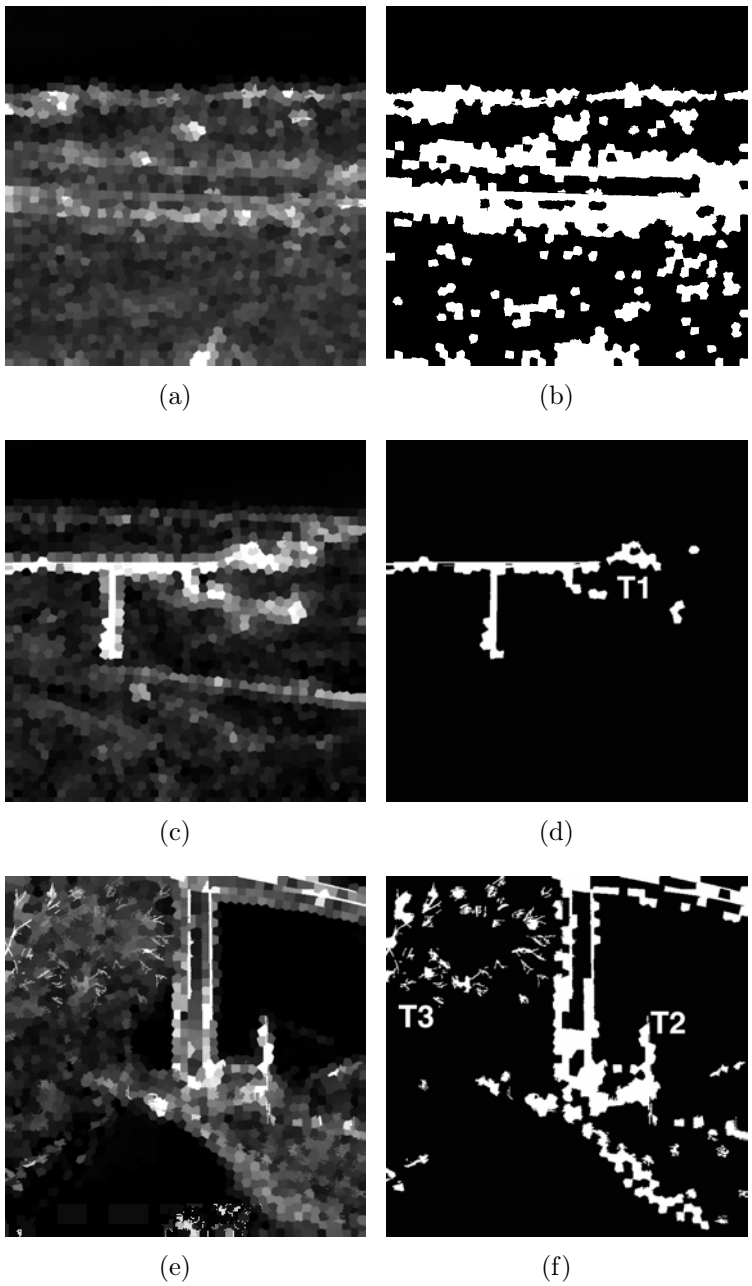


Figure 3.16: Change detection results with Achanta segmentation: (a), (c), (e) change matrices; (b), (d), (f) final change maps.

in a similar way we did in Section 3.7, to generate the final change detection result. Figures 3.17(a), (d) and (g) present the results for the three image pairs shown in Figure 3.12 by using the change detection method based on image differencing. We can observe that this method is very sensitive to shadows: the effect of shadow makes the region "C1" near the bridge of Figure 3.17(d) be wrongly labeled as changed. And the region of sky "C2" in Figure 3.17(d) belonging to unchanged area is completely detected as changed.

Therefore, this comparative method used in our observatories shows the worse performance as described in Table 3.3. It produces a mean overall accuracy of 67.86%, mean F1 score of 16.18%, mean omission error of 52.27% and 28.34% in changed region and unchanged region, respectively, mean commission error of 85.54% in changed region and 8.86% in unchanged region.

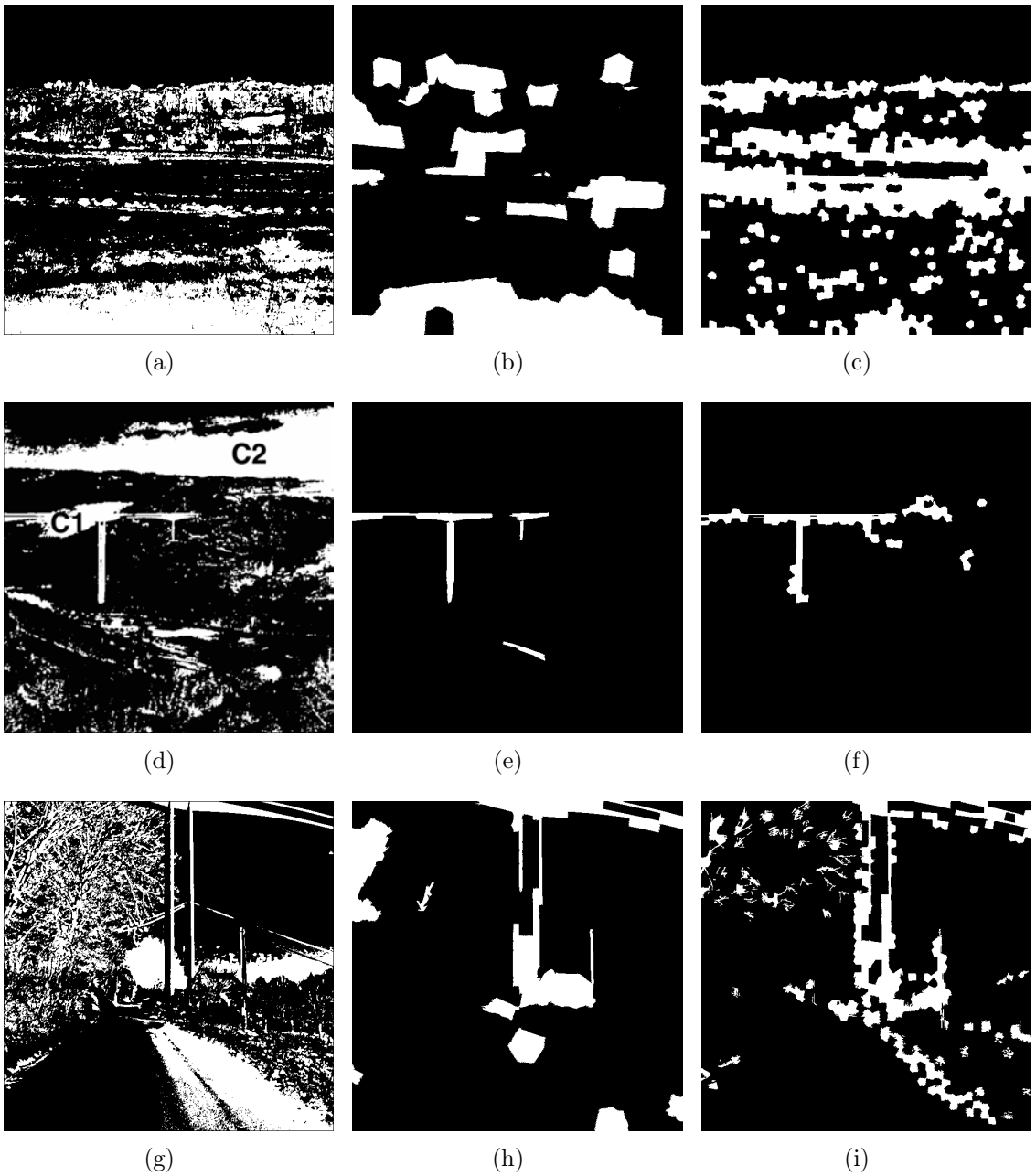
Table 3.3: Performance evaluation with image differencing change detection

	Overall accuracy	Omission error		Commission error		F1 score
		changed	unchanged	changed	unchanged	
Mean [std]	67.86% [8.01%]	52.27% [22.07%]	28.34% [8.80%]	85.54% [23.10%]	8.86% [11.56%]	16.18% [16.54%]
Median	68.92%	51.02%	28.41%	93.65%	2.27%	11.44%

3.9 Conclusion

In this chapter, we introduced an automatic change detection approach to detect the changed and unchanged pixels. Compared to multi-channel remote sensing images, our image collection taken by digital cameras usually suffers from the lack of information at the pixel level. Therefore, we detect changes in groups of pixels instead of individual pixels, i.e., we perform the tests on superpixels because of their ability to aggregate local information. Textural appearance is also a useful property for change detection of landscape images. We refer to texton histogram and gray-level mean to construct the feature set for each superpixel of each image pair. Finally, we measure the Mahalanobis distance value between the two feature sets of the superpixel pair to produce the change matrix. The binary change map is obtained by thresholding the change matrix as explained in Section 3.7.

We compared two superpixel segmentation algorithms: *Shi segmentation* and *Achanta segmentation*. The segmentation results show that:



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 3.17: Comparative change detection results obtained using (a), (d), (g) image differencing, (b), (e), (h) Shi segmentation and (c), (f), (i) Achanta segmentation.

- Shi segmentation is able to generate compact superpixels with a regular shape at the price of frequent boundary adherence errors and a rather long runtime.
- Achanta segmentation can provide compact superpixels with a better boundary adherence and is faster and more memory efficient.

Obviously, Achanta method has turned out to be more suitable for our application of automatic change detection.

We knew that the problem of change detection in photographic landscape observatories was challenging in particular because of several issues like weather or time of the day inducing changes in the images as strong as the significant changes we are looking for. The automatic method we propose partly solve the problem giving better performance than a simple image differencing. But one can notice that the F1 scores we obtain for the tested image pairs are the best achievable with a global thresholding strategy to obtain the final binary change maps. However, the results we obtain do not meet the requirements for an operational implementation into a photographic landscape observatory.

Therefore, in order to improve the performance and to take into account the intended application, we propose in the next chapter a semi-automatic approach in which the user is requested to provide clues to initiate the process by giving samples of changed and unchanged pixels.

Chapter 4

Interactive Change Detection based on Dissimilarity Image and Supervised Classification

Abstract of the chapter

In this chapter, we investigate a new approach of change detection between two images of the same landscape acquired at different dates. In order to detect only changes that are significant for a specific application, the user is involved at the beginning of the process.

We propose an interactive change detection method in which the user gives sample pixels belonging to the changed and unchanged objects. Each pixel is described by a feature vector. Then, a classifier trained with this data is used to label all the pixels of the image pair. Finally, a post-processing step using basic binary morphology operators is used to correct some classification errors.

The data used for the training step and the binary classification consist of features describing each chronological pixel pair. We use 16 features mainly obtained from the dissimilarity between the gray-levels of the neighbors of the pixel pair.

We evaluate the performance of the proposed method on the data set already used for the automatic method method presented in Chapter 3. We study the behavior of two classification tools: decision trees and random forests. We also evaluate the influence of the neighborhood size considered in the dissimilarity computation.

Compared to the automatic one, the proposed method achieves better performance. Although slower, random forests give better results than decision trees. Satisfactory results are obtained using a small neighborhood size for computing the dissimilarity image. The experiments we carry out show the potential of the proposed approach.

4.1 Introduction

The significance of a change between two images relies on the underlying application. For example, if the application deals with long term impact on natural landscapes of the road infrastructure, the apparition of the bridge in the image pair of Figure 4.1 is significant, but marks left by earth-moving machines are not. On the opposite, if the application concerns the study of consequences of civil engineering works, then areas where vegetation has disappeared because of excavation around the bridge pillars are significant.

(a) Image 1 (b) Image 2

Figure 4.1: One image pair illustrating the significance of changes.

Without any user intervention, a fully automatic change detection method is not able to give an application-dependent interpretation of change significance. So knowledge has to be taken into account by the change detection system. One solution may be to use a machine learning approach and, in an early phase, to train the system with samples corresponding to different types of application. But this would require to label a large amount of data.

The solution we propose consists in requiring the user to interact whenever he wants to process an image pair. This interaction is reduced to the selection of pixels belonging to the two classes: changed and unchanged. These samples are used to train a classifier which then classifies all the other pixels.

In order to obtain a discriminant description of the pixels, it is useful to take also into account the neighbors of each pixel. This is a classical approach not only used for change detection, but also, for example, in local matching methods used in binocular stereo vision [CC11]. We propose to describe each pixel by a feature vector mainly based on the dissimilarity image. These vectors are the data used to train and to exploit a classifier for obtaining a binary change map.

The main steps of the proposed method are summarized in the diagram on Figure 4.2

whose details are given below.

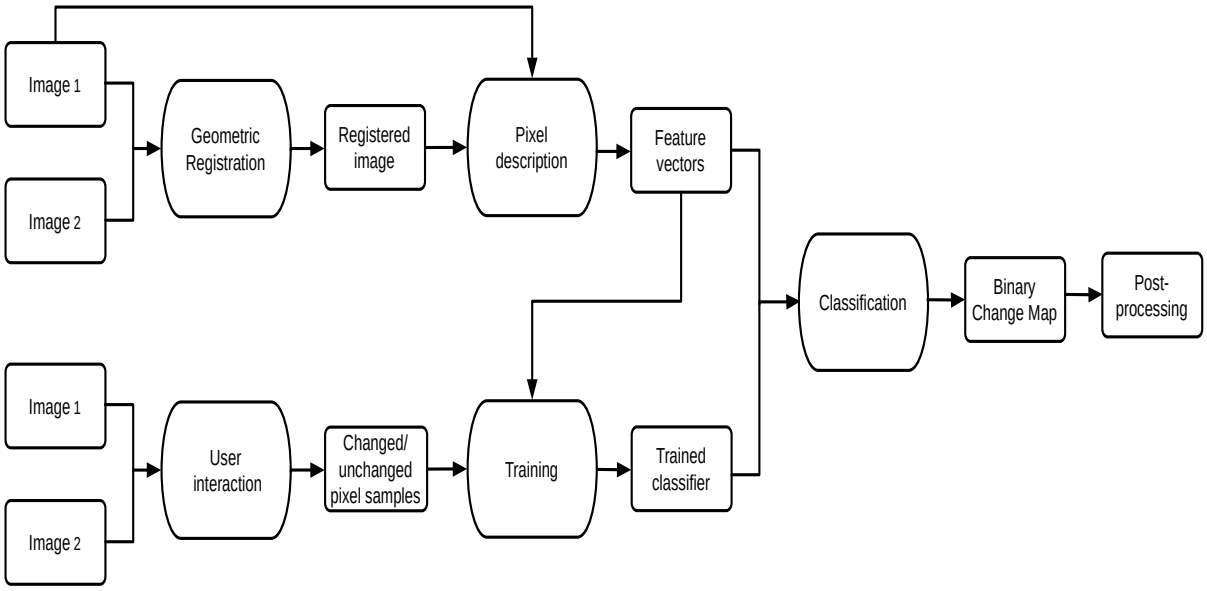


Figure 4.2: Flow chart of the learning step of the interactive change detection method.

4.2 Description of Pixels

In order to be classified as changed or unchanged, each pixel needs to be described by numerical attributes. A first step consists in considering not only the gray-level of the pixel but also the gray-levels of the neighbors surrounding the pixel. For a square neighborhood of size $N \times N$, each pixel is described by a vector of $n = N^2$ gray-levels.

The simple difference of two gray-levels calculated in the image differencing technique is then replaced by the comparison of two gray-level vectors. Many measures of similarity or dissimilarity of two gray-level vectors have been proposed, particularly for pixel matching in binocular stereo vision applications [CC11]. We propose to use the simple sum of squared differences measure which is the square of the Euclidean distance and is defined by :

$$\mathrm{SSD}(x, y) = \|\mathbf{f}_1 - \mathbf{f}_2\|^2 \tag{4.1}$$

where $\mathbf{f}_1$ and $\mathbf{f}_2$ are the vectors of the gray-levels of the pixels in the $N \times N$ window centered at pixel (x, y) in the first and second date image, respectively.

Dissimilarity measure SSD gives high values as the two neighborhoods differ. The matrix of all SSD values can be linearly transformed to a gray-level image that we call the dissimilarity image. An example is given in Figure 4.3(c).

(a) Image 1 (b) Image 2 (c) SSD

Figure 4.3: Dissimilarity score visualization.

From the dissimilarity image, we extract texture features based on the co-occurrence matrices. Gray-level co-occurrence matrices [HSD73] are a classical tool for extracting textural features, which describes the frequency of one gray-level occurring in a specified spatial relationship to another gray-level within the area considered. For a given image I with size of $n \times m$, we can calculate a co-occurrence matrix with an offset $(\Delta x, \Delta y)$:

$$C_{\Delta x,\Delta y}(i,j) = \sum_{x=1}^{n}\sum_{y=1}^{m} \begin{cases} 1, & \text{if } I(x,y) = i \text{ and } I(x+\Delta x, y+\Delta y) = j \\ 0, & \text{otherwise} \end{cases}, \tag{4.2}$$

where i and j are the image intensity values of image I and x and y are the spatial positions in image I.

We measure the four important properties [ST99] of the co-occurrence matrices of the dissimilarity image:

$$\text{Energy: } f_1 = \sum_i \sum_j p(i,j)^2, \tag{4.3}$$

$$\text{Contrast: } f_2 = \sum_i \sum_j (i-j)^2 p(i,j), \tag{4.4}$$

$$\text{Correlation: } f_3 = \frac{\sum_i \sum_j (i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}, \tag{4.5}$$

$$\text{Homogeneity: } f_4 = \frac{p(i,j)}{1+|i-j|}, \tag{4.6}$$

where $p(i,j)$ refers to the (i,j)th normalized entry a the co-occurrence matrix.

We then calculate mean and standard deviation of each property for each pixel, which gives an 8-dimensional vector of co-occurrence properties for each pixel.

We propose to add to these texture features the local mean, standard deviation and entropy of the dissimilarity image. In a window of size $N \times N$, the entropy is defined by:

$$H = -\sum_{k=0}^{K-1} P_k \log_2 P_k \tag{4.7}$$

where $P_k = n_k/n$ is the probability of gray-level k, n_k is the number of pixels of gray-level k , $n = N^2$ is the number of pixels and K is the number of gray-levels.

In addition to features calculated from the SSD dissimilarity image, we describe the relation between the two-date gray-level vectors by the mean of linear correlation. Like Im and Jensen [IJ05] proposed, we calculate the slope and the intercept of the linear regression line between the two gray-level vectors. An example of such a regression line is shown in Figure 4.4.

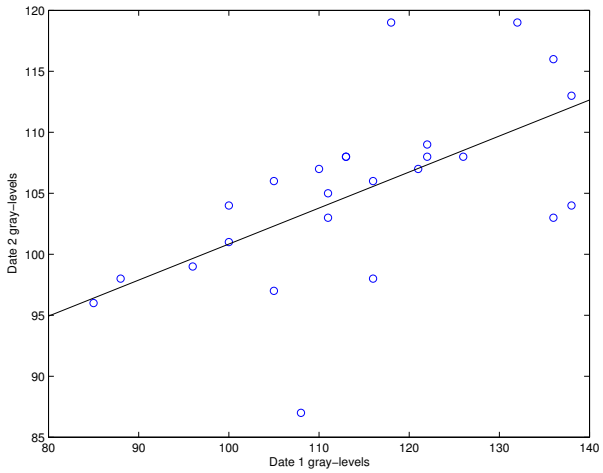


Figure 4.4: Regression line from the gray-levels of the two-date neighborhoods.

More precisely, slope and intercept are given by:

$$H_{slope} = \frac{\text{COV}_{12}}{{s_1}^2}, \qquad H_{intercept} = m_1 - H_{slope} m_2, \tag{4.8}$$

where m_1 and m_2 are the means of $\mathbf{f}_1$ and $\mathbf{f}_2$ values, and s_1 and s_2 are the standard deviations of $\mathbf{f}_1$ and $\mathbf{f}_2$ values, respectively. Covariance of $\mathbf{f}_1$ and $\mathbf{f}_2$ is given by:

$$\text{cov}_{12} = \frac{1}{n-1}\left(\sum_{k=1}^{n}(f_1(k) - m_1)(f_2(k) - m_2)\right), \tag{4.9}$$

where $f_1(k)$ and $f_2(k)$ are the k^{th} values in vectors $\mathbf{f}_1$ and $\mathbf{f}_2$, respectively.

Each pixel is described by the following 16 features:

- gray-level in the first image;
- gray-level in the second image;
- dissimilarity value;
- local mean of the dissimilarity image;
- local standard deviation of the dissimilarity image;
- local entropy of the dissimilarity image;
- local mean of the co-occurrence energy;
- local standard deviation of the co-occurrence energy;
- local mean of the co-occurrence contrast;
- local standard deviation of the co-occurrence contrast;
- local mean of the co-occurrence correlation;
- local standard deviation of the co-occurrence correlation;
- local mean of the co-occurrence homogeneity;
- local standard deviation of the co-occurrence homogeneity;
- slope of the regression line;
- intercept of the regression line.

4.3 Simple Interactive Sample Region Extraction

The idea of interactive binary segmentation consists in asking a user to designate some pixels in an image belonging to the background and others to the foreground. Then, this data is used to train a classifier in order to assign a label to each pixel in the image. We apply the same idea on the two classes: changed and unchanged pixels. These samples are then used as input data of a training algorithm. Figure 4.5 shows the user-defined sample regions, where the red and black colors denote the unchanged and changed samples, respectively.

4.4 Supervised Classification

Once the feature set is defined for training data, a proper classifier should be selected to perform the supervised classification involved in change detection. We experimented two classifiers: decision trees and random forests.

4.4.1 Decision Trees

Decision tree learning [SL91] is a supervised algorithm. An important property of decision trees is their semantic interpretation that is a natural simulation of human learning.

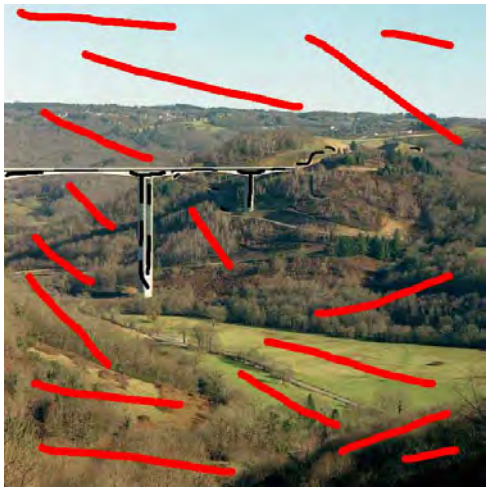


Figure 4.5: User-defined sample regions.

Compared with other machine learning algorithms, decision tree learning is simpler to use and can be robust to incomplete and noisy data due to post-pruning techniques [Qui86, PM01]. These advantages make decision trees perform well in data classification [Qui86, BFSO84, PM01, PT98, BA97, Mit97].

A decision tree is a rooted tree, which contains zero or more internal nodes representing attributes (all nodes except the root and the leaves), and one or more leaf nodes representing a class (terminal nodes with no children), branches representing the possible attribute values. If the children of each node are ordered, i.e., normally from left to right, then it will be called an ordered tree. A binary tree is an ordered tree in which each child of a node is separated either as a left child or a right child, and no node has more than one left child or more than one right child, that is, the root node and all internal nodes have two child nodes. All non-terminal nodes contain splits.

To classify a data, the decision tree is traversed from top to bottom by testing the data attributes at each internal node that comes in the way until a leaf node is encountered. The output is a class label. The number of the classes is finite and their values are defined beforehand.

Algorithms to build a decision tree (usually called tree induction, tree building or tree growing) take a training data set as input, each of which is completely described by a set of attributes and a class label. Several methods have been presented to construct a decision tree, such as C4.5 [Qui93], ID3 (i.e., Iterative Dichotomiser 3) [Qui86], and CART (i.e., Classification And Regression Tree) [BFSO84]. These methods generally use the recursive-partitioning algorithm, and their inputs require a set of training examples, a splitting rule, and a stopping rule.

Due to the noise (class mislabelling or attribute mislabelling), poor ability of leaf nodes

to represent instances, and a missing class label or attribute, a tree may casually expand. Such a redundant decision tree generally leads to misclassifications, thus, it is necessary to build a simple decision tree by minimizing the number of nodes or controlling the depth [BA97]. The commonly used solutions are pre-pruning and post-pruning. Pre-pruning techniques are computationally efficient, since they can keep from the expansion of tree by setting the terminal criterion before the use of complete training data set for the building process [Qui86]. However, the pre-pruning techniques might terminate the tree building prematurely, which results in inconsistent classification performance. This makes post-pruning techniques more widely used. Post-pruning techniques simplify the tree by removing or replacing unimportant branches and unknown outputs after a decision tree has been generated by using all the training data.

C4.5 algorithm using post-pruning technique to optimize a tree and handing both continuous and discrete attributes has been proved to be one of the best tools to build a decision tree. We use it in our method to train the classifier.

4.4.2 Random Forests

Random forests are a combination of decision trees such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest [Bre01]. Breiman [Bre01] defined a random forest as a classifier consisting of a collection of tree-structured classifiers whose each tree casts a vote for a class and the forest chooses the most popular class.

Random forest are an efficient tool for classification that is robust against overfitting. It can provide very high performance of classification by filtering out a kind of randomness. These properties therefore encourage us to use this classifier.

4.5 Binary Mathematical Morphology

Mathematical morphology [Ser82] is a theory for the analysis and description of geometrical structures based on set theory. In image processing, mathematical morphology is used to deal with the interaction between an image and a certain structuring element, such as using the basic morphology operators: erosion and dilation.

Normally, the structuring elements are small compared to the image. When processing digital images, simple binary structuring elements like a cross or a square are typically used, such as the ones shown in Figure 4.6.

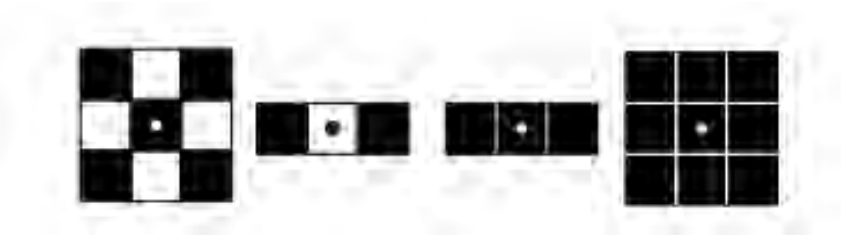


Figure 4.6: Examples of mathematic morphology structuring elements: the central pixel denotes the origin, black pixels are the components of structuring elements.

4.5.1 Dilation and Erosion

Dilation and erosion are the basic operations in mathematical morphology. As in binary morphology, let A be a binary image regarded as a subset of a Euclidean space or the integer grid, and B be a structuring element. Then, the dilation operation of A by B is defined as:

$$A \oplus B = \{z \in E \mid (B^s)_z \cap A \neq \varnothing\}, \tag{4.10}$$

where E is a Euclidean space or the integer grid, B^s denotes the symmetric of B, and

$$B^s = \{x \in E \mid -x \in B\}. \tag{4.11}$$

Dilation which is a shift-invariant operator is used to fill the small holes or expand the shapes contained in the input image.

In the opposite, the erosion operation is used to reduce the shapes in the image, which is defined as:

$$A \ominus B = \{z \in E \mid (B)_z \subseteq A\}, \tag{4.12}$$

where $(B)_z$ is the translation of B by the vector z, such as,

$$(B)_z = \{b + z \mid b \in B\}, \forall z \in E. \tag{4.13}$$

4.5.2 Opening and Closing

The composition of the erosion and dilation operations can generate the opening and closing operations.

Morphological opening is the dilation of the erosion of a set A by a structuring element B, which can be calculated as follows:

$$A \circ B = \{(A \ominus B) \oplus B\}. \tag{4.14}$$

The opening operation is used to remove unwanted structures in an input image, e.g.,

noise.

Morphological closing is the erosion of the dilation of a set A by a structuring element B, which can be calculated as follows:

$$A \bullet B = \{(A \oplus B) \ominus B\}. \tag{4.15}$$

The closing operation is used to merge or fill the structures in an input image.

The output of the classifier includes misclassified small regions and needs a post-processing step. We propose to simply apply closing followed by opening. The closing operation allows to remove false negatives, while opening can remove false positives.

4.6 Experiments and Discussions

From a user point of view, our proposed interactive change detection method is divided into two main steps: interaction and classification. The user designates with the mouse samples of changed and unchanged pixels. Then, the program classifies the remaining pixels.

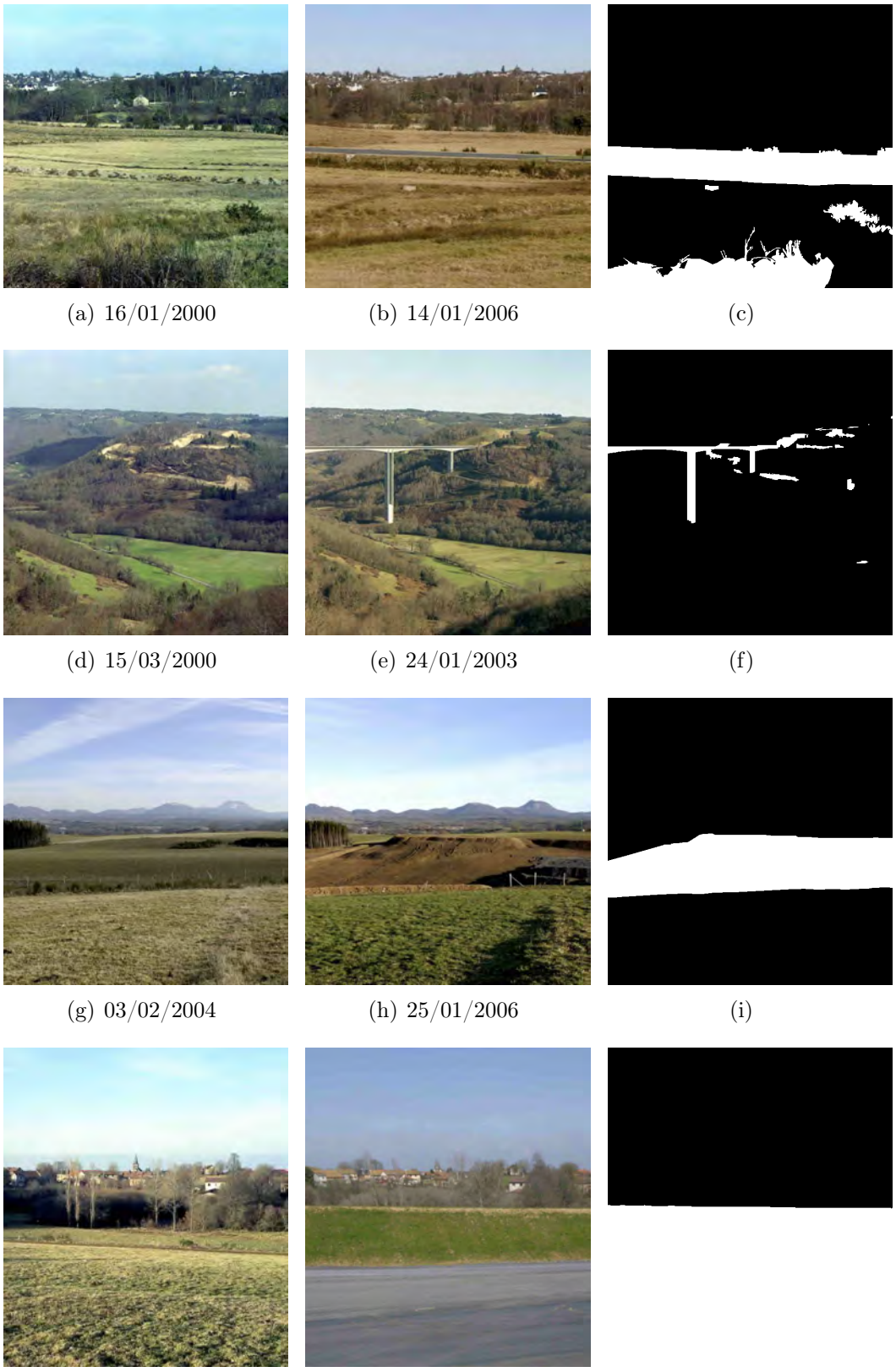
We tested the method on the same data set we used for evaluating the automatic method. It consists of 19 images corresponding to 7 different scenes, all images are grouped into 12 image pairs. Each image pair consists of one image before the construction of the highway and the other one during the construction. A part of the image series was shown in the previous chapter (Figures 3.10 and 3.11).

In order to measure the performance of the method, like in the previous chapter, we compared the results to the ground truth. Four image pairs and the associated reference change maps are given in Figure 4.7. These pairs will be used in the following sections for visually illustrating the results we obtained.

For each image pair, we designated around 20000 sample pixels accounting for around 4% of all pixels. Examples of sample pixels selected by the user are shown in Figure 4.8 where black pixels belong to changed objects and red pixels designate unchanged pixel samples. Among these pixels, 10% belong to changed regions. These samples were used to train a classifier which then predicts the labels of the 96% remaining pixels and produces a binary change map. Finally, this map was post-processed by closing and then opening using the 3×3 structuring element.

For the training and the classification steps, each pixel is described by a 16-dimensional feature vector as explained in Section 4.2. When calculating entropy, we used a 5×5 window size for all dissimilarity images. We computed the co-occurrence matrices in 3×3 windows, using 9 gray-levels and $[0\ 1; -1\ 1; -1\ 0; -1\ -1]$ as offsets.

We describe below the results obtained using two types of classifiers: decision trees and



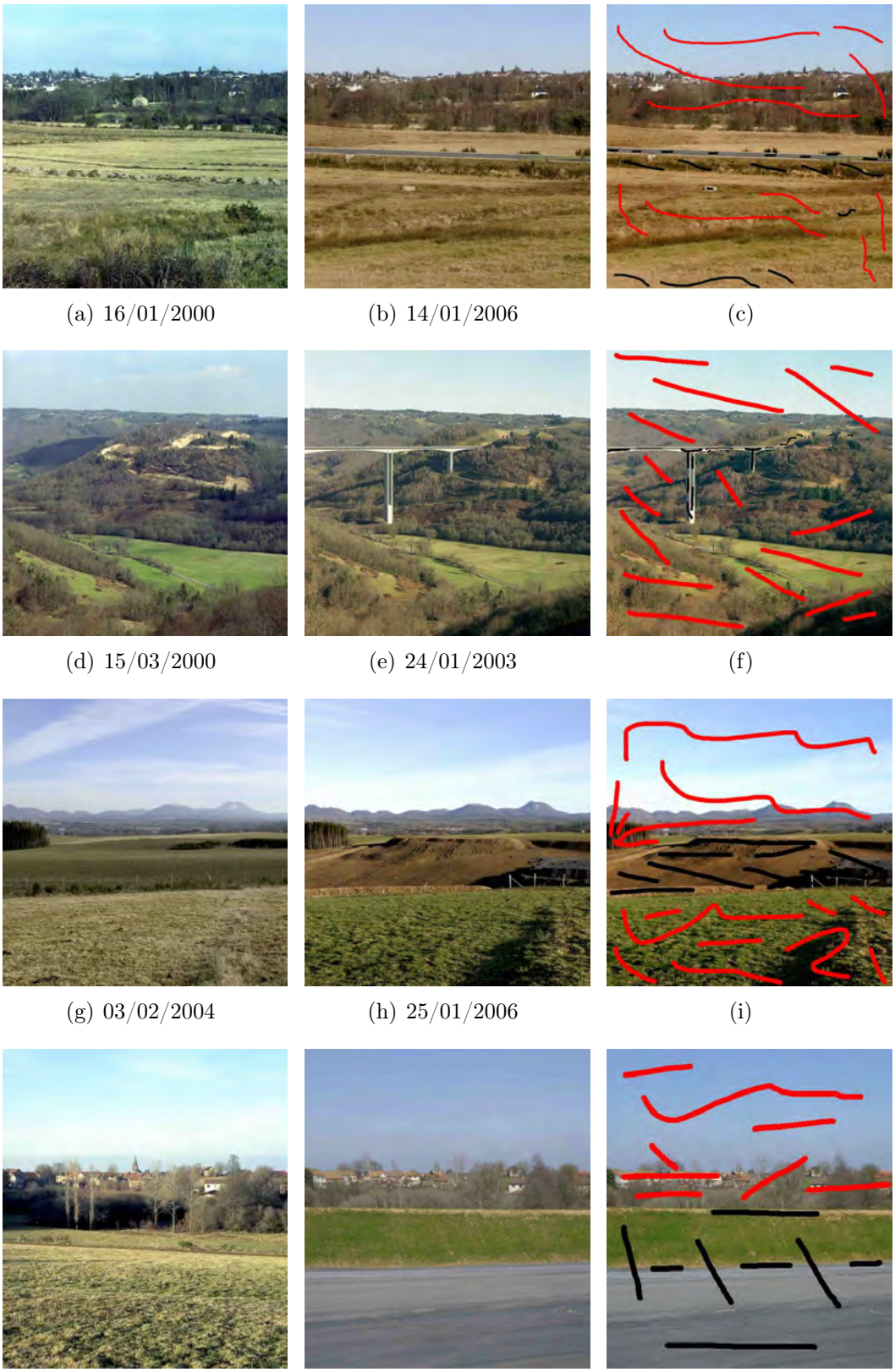
(a) 16/01/2000 (b) 14/01/2006 (c)

(d) 15/03/2000 (e) 24/01/2003 (f)

(g) 03/02/2004 (h) 25/01/2006 (i)

(j) 20/02/2000 (k) 01/04/2005 (l)

Figure 4.7: Examples of image pairs with ground truth: (a), (d), (g), (j) first and (b), (e), (h), (k) second date images; (c), (f), (i), (l) ground truth change maps.



(a) 16/01/2000 (b) 14/01/2006 (c)

(d) 15/03/2000 (e) 24/01/2003 (f)

(g) 03/02/2004 (h) 25/01/2006 (i)

(j) 20/02/2000 (k) 01/04/2005 (l)

Figure 4.8: Examples of sample pixels: (a), (d), (g), (j) first and (b), (e), (h), (k) second date images; (c), (f), (i), (l) sample pixels selected by the user (black for changed and red for unchanged pixels).

random forests. For each classifier, we evaluated the influence of the size of the window used to calculate the dissimilarity images. We report here the results obtained with the sizes: 3×3, 5×5 and 7×7.

4.6.1 Experimental Results using Decision Trees

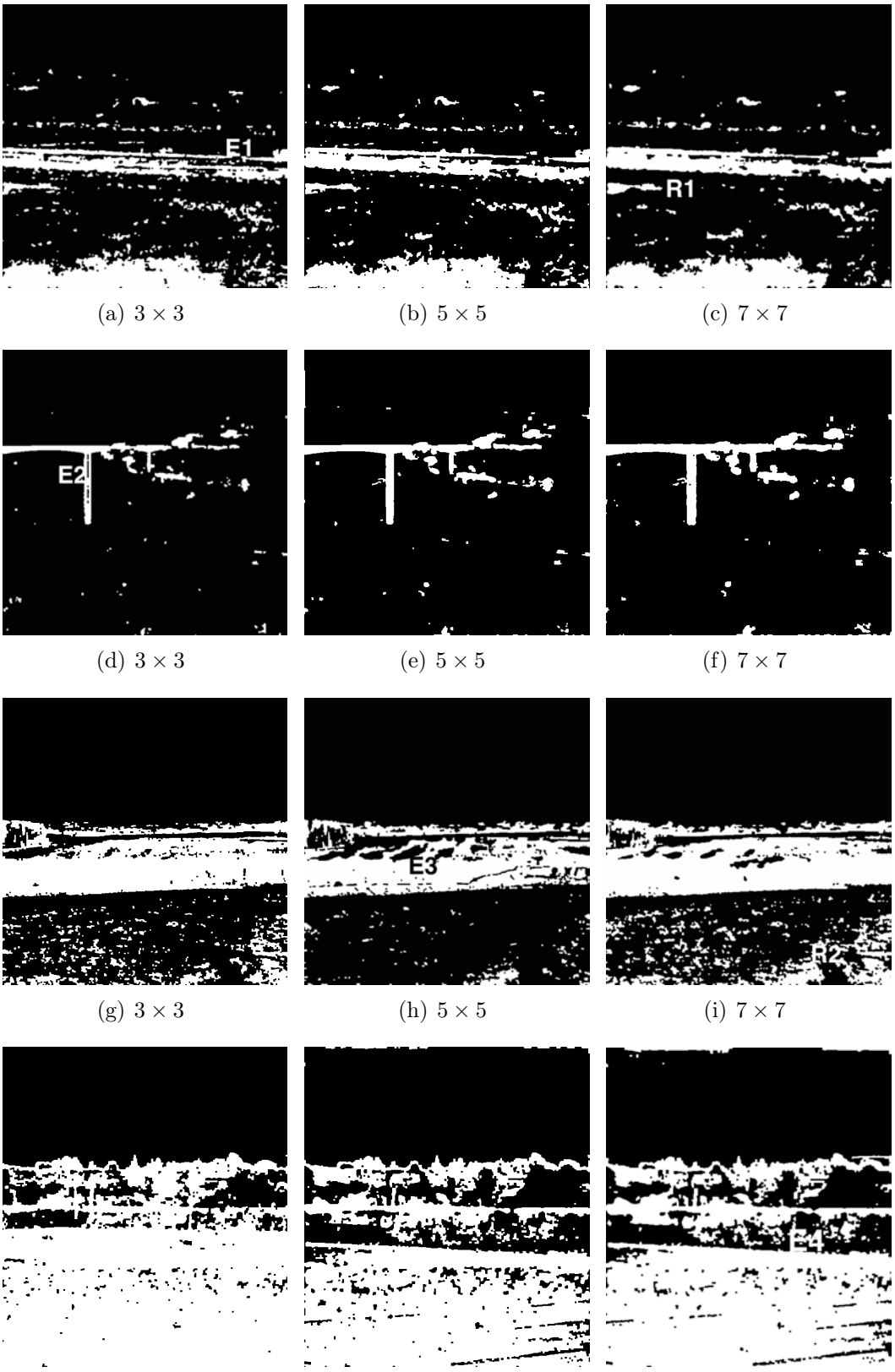
We present here some results obtained using a decision tree. Figure 4.7 presents four of the 12 image pairs we used and the associated ground truth change maps.

Binary change maps, obtained using 3×3, 5×5 and 7×7 neighborhood sizes for computing the SSD dissimilarity images, are given in Figure 4.9. One can observe that changed areas are obtained relatively completely, but some details are not satisfactorily detected, such as "E1", "E2", "E3" and "E4" in Figure 4.9(a), (d), (h) and (l), respectively. In addition, some unchanged areas are wrongly detected as changed such as "R1" and "R2" in Figure 4.9(c) and (i). These errors can be explained by the presence of shadows generating false positives and by the fact that the variation of texture is sometimes similar for change and unchanged areas.

From the visualization of change detection results in Figure 4.9, we find that different neighborhood sizes used to compute SSD dissimilarity images lead to different results. A larger neighborhood size generates more regular change maps due to the smoothing effect by reducing high frequency components in the dissimilarity image, while more often wrongly detects changed and unchanged pixels. A smaller neighborhood size keeps detailed change information, but is sometimes sensitive to the effect of noise. Nevertheless, it is quite difficult to visually interpret the produced change maps that give only a subjective evaluation of the performance.

Thus, in Table 4.1, we present the values obtained for the same performance criterions used in Chapter 3. For the three neighborhood sizes, 3×3, 5×5, and 7×7, we give the mean, the standard deviation and the median of the values of the overall accuracy, the omission errors, the commission errors and the F1 score. We can see that the 3×3 size achieves the best "global performance" with a mean overall accuracy near 95% and a F1 score near 67%. This global performance decreases when the neighborhood size increases. However, for few specific performance measures, bigger neighborhood sizes can outperform the 3×3 size, for example the 5×5 neighborhood size for the commission error for changed pixels.

It can be noticed that standard deviation values of some measures indicate that they are rather dispersed. It is, for example, the case of F1 scores and commission errors of changed pixels for which standard deviations are greater than 10% and even 13% for omission errors of changed pixels. This dispersion shows that the performance of the method varies depending on the image pair. Moreover, in most cases, the median values



(a) 3×3 (b) 5×5 (c) 7×7

(d) 3×3 (e) 5×5 (f) 7×7

(g) 3×3 (h) 5×5 (i) 7×7

(j) 3×3 (k) 5×5 (l) 7×7

Figure 4.9: Change detection results obtained from image pairs of Figure 4.7 using a decision tree: the size of the neighborhood used for computing the dissimilarity image is given under each binary change map.

differ from the mean values. These differences show that the performance for few image pairs is largely below the performance for the other images. For example, in the third image pair (g) and (h) in Figure 4.7, small temporal variation of texture corresponds both to changed and unchanged areas and makes difficult for the classifier to distinguish correctly between the two classes.

Another problem can be reported for the fourth image pair (j) and (k) in Figure 4.7 for which, contrary to the other pairs, registration errors are not negligible and influence the change detection result. Our geometric image registration procedure, described in Section 3.2, involves the estimation of the homography parameters from correspondences of points of interest. Unfortunately, due to important changes in the foreground and the surface occupied by the sky, point correspondences have been found only in a narrow band in the center of the images. This bad spatial distribution causes errors in the estimation of the homography parameters. An even more serious problem is that the homography model we used for registration is based on the assumption of a common position of the camera center, i.e., only a rotation about this center is allowed between the two acquisitions. But, for this pair, the initial position of the tripod has probably not been located accurately when the second image has been acquired. Indeed, the mark on the ground made at the first acquisition has been probably removed during the construction of the roadway.

Regarding the processing time, because we used a pure Matlab implementation, it is not fast enough to obtain a comfortable tool. The construction of the decision tree from sample pixels given by the user takes around 15 seconds and the classification of the remaining pixels roughly the same time.

4.6.2 Experimental Results using Random Forests

We made the same experiments as those described in the previous chapter, replacing decision trees by random forests. The results are illustrated by Figure 4.10 and Table 4.2.

We used 30 trees per forest and we set the percentage of training data for each tree to 33%.

The examples of change maps show that the use of random forests instead of decision trees improves the change detection performance. For example, some changed pixels are now correctly detected like in the bush “E5” in Figure 4.10(b) and on the bridge pier “E6” in Figure 4.10(d).

Performance criterion values given in Table 4.2 confirm this improvement. The overall accuracy benefits from a small improvement, but the mean F1 score raise from 67% to more than 73% for the 3×3 neighborhood size for computing the SSD dissimilarity images.

Table 4.1: Performance evaluation of the interactive method with a decision tree.

	Overall accuracy	Omission error		Commission error		F1 score
		changed	unchanged	changed	unchanged	
3×3						
Mean [std]	94.99% [4.30%]	27.46% [13.69%]	4.12% [4.82%]	36.91% [13.16%]	2.47% [2.91%]	66.92% [11.98%]
Median	97.82%	25.01%	1.35%	35.62%	1.18%	69.56%
5×5						
Mean [std]	94.56% [5.27%]	31.76% [15.93%]	3.41% [3.63%]	35.08% [11.95%]	3.87% [6.09%]	64.85% [10.56%]
Median	97.30%	25.63%	2.21%	36.08%	1.03%	67.00%
7×7						
Mean [std]	93.93% [5.18%]	28.01% [16.52%]	4.66% [4.61%]	42.37% [14.46%]	3.39% [5.54%]	62.16% [11.47%]
Median	96.26%	22.74%	3.16%	45.52%	1.13%	62.80%

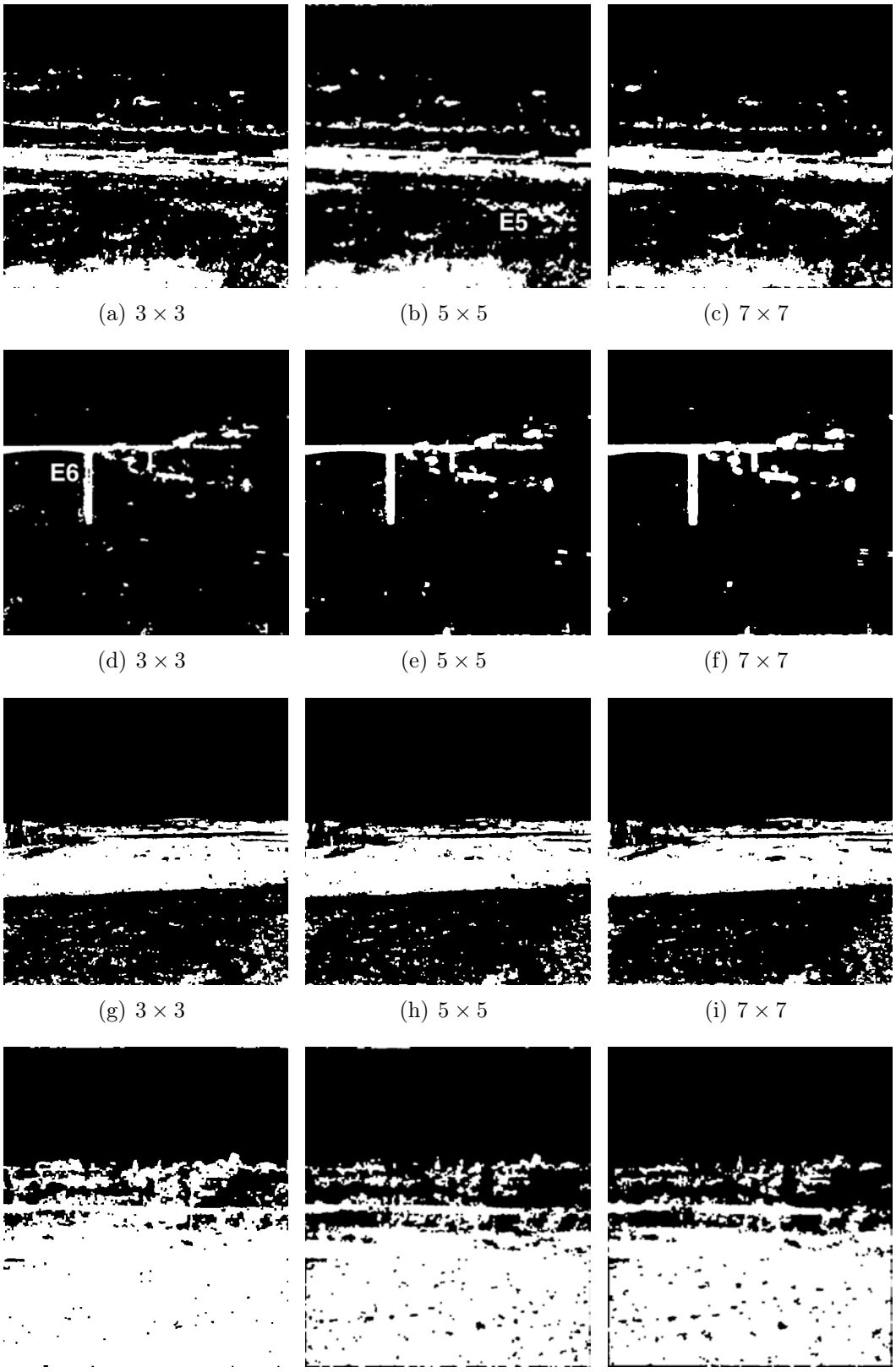
The omission error for changed pixels is noticeably improved, decreasing of almost 30% to around 20% for the three tested neighborhood sizes. A slight improvement can also be observed for the commission error that is reduced from 37% to almost 30% for the 3×3 neighborhood size. The dispersion of all the performance criterion values is also reduced as shown by the standard deviation values given in table 4.2.

We can also notice that, in maps 4.10(j) and (k), the registration errors of the corresponding image pair, already mentioned in Section 4.6.1, bring false negatives on the border of the change maps.

Regarding the influence of the neighborhood size, the same observations as for decision trees can be made. A small size appears to give the best results.

Regarding the processing time, the performance improvement comes at the price of a slower training and a slower classification. The training step needs about 5 minutes and the classification of the remaining pixels of the image pair takes around 9 minutes. The implementation we used is partly written in Matlab, but mainly in C++ (ALGLIB[1]).

[1] ALGLIB can be found at `http://www.alglib.net/`.



(a) 3×3 (b) 5×5 (c) 7×7

(d) 3×3 (e) 5×5 (f) 7×7

(g) 3×3 (h) 5×5 (i) 7×7

(j) 3×3 (k) 5×5 (l) 7×7

Figure 4.10: Change detection results obtained from image pairs of Figure 4.7 using a random forest: the size of the neighborhood used for computing the dissimilarity image is given under each binary change map.

Table 4.2: Performance evaluation of the interactive method with a random forest.

	Overall accuracy	Omission error		Commission error		F1 score
		changed	unchanged	changed	unchanged	
3×3						
Mean [std]	95.80% [3.83%]	20.05% [12.75%]	3.72% [4.12%]	30.57% [9.32%]	1.73% [2.04%]	73.59% [9.02%]
Median	97.98%	18.00%	1.62%	29.85%	0.65%	74.52%
5×5						
Mean [std]	95.53% [3.74%]	19.57% [10.33%]	3.45% [3.03%]	34.37% [12.05%]	2.29% [3.45%]	71.47% [8.74%]
Median	97.5%	17.61%	2.17%	35.97%	0.59%	70.96%
7×7						
Mean [std]	95.39% [3.57%]	19.39% [10.22%]	3.47% [2.76%]	37.31% [14.08%]	2.39% [3.68%]	69.59% [9.87%]
Median	97.11%	17.04%	2.58%	41.78%	0.58%	68.05%

4.6.3 Comparative Experiment

In this section, we compare the performance obtained by the automatic method we described in Chapter 3 and the interactive method considered in the current chapter.

For the automatic method, we use Achanta segmentation because it gives the better results in our experiments. For the interactive method, we use the random forest variant and the 3×3 neighborhood size for the computation of the SSD dissimilarity image for the same reason.

The behavior of the two methods is illustrated by Figure 4.11 whose rows correspond to the four examples of image pairs we use since the beginning of the chapter. The first column shows the change maps obtained using the automatic method. The second column presents the results given by the interactive method. Finally, we show again the ground truth in the last column.

Then Table 4.3 contains performance criterion values extracted from Table 3.2 for the automatic method and Table 4.2 for the interactive method.

Figure 4.11 shows the obvious sensitivity of the method to the segmentation results. Indeed, when a superpixel is misclassified, all the pixels belonging to this superpixel are

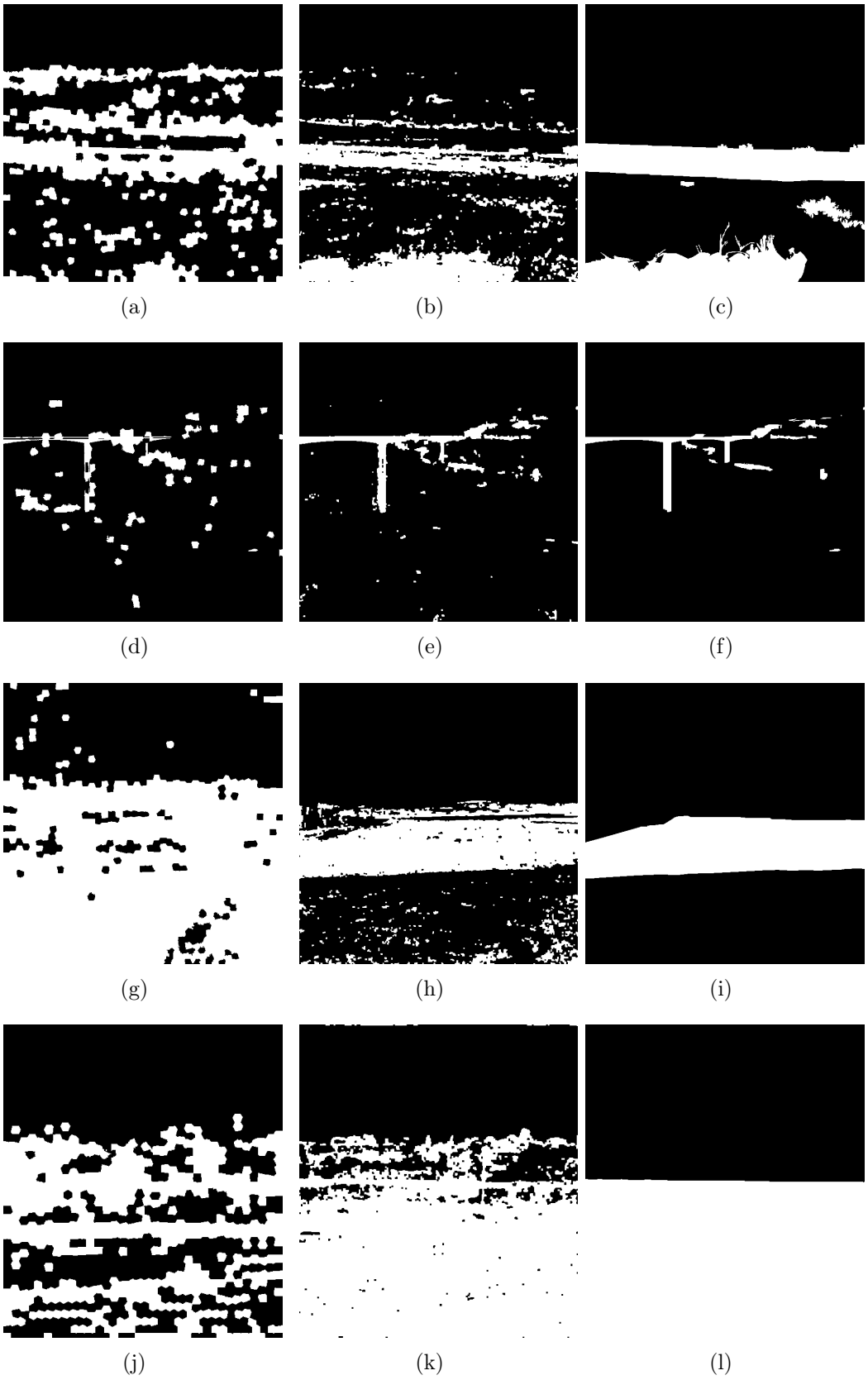


Figure 4.11: Comparative change detection results: (a), (d), (g) and (j) automatic method with Achanta segmentation; (b), (e), (h) and (k) interactive method with random forests using 3×3 neighborhood size for SSD dissimilarity image; (c), (f), (i) and (l) ground truth.

Table 4.3: Comparison of performance measures between the proposed automatic and interactive methods.

	Overall accuracy	Omission error		Commission error		F1 score
		changed	unchanged	changed	unchanged	
Automatic method using Achanta segmentation						
Mean [std]	82.83% [21.57%]	43.89% [22.61%]	16.08% [24.26%]	62.54% [19.19%]	5.63% [10.39%]	41.40% [15.17%]
Median	95.50%	46.18%	3.25%	63.57%	1.70%	44.70%
Interactive method using random forests (3×3)						
Mean [std]	95.80% [3.83%]	20.05% [12.75%]	3.72% [4.12%]	30.57% [9.32%]	1.73% [2.04%]	73.59% [9.02%]
Median	97.98%	18.00%	1.62%	29.85%	0.65%	74.52%

misclassified. In addition, if a superpixel straddles two different objects, it may include changed and unchanged pixels. In this case, misclassified pixels are unavoidable.

The superior performance of the interactive method is confirmed by the values reported in Table 4.3. All the criterions show a clear improvement. The mean overall accuracy is increased by 13%. The mean F1 score is increased from 41% to 74%. The improvement of omission errors and commission errors is much more noticeable. Mean omission error is divided by two for changed pixels and by four for unchanged pixels. Mean commission error is divided by two and by three for changed and unchanged pixels respectively. Dispersion of the values of all the criterion is also clearly reduced, especially for the overall accuracy and the omission and commission errors of the unchanged pixels.

4.7 Conclusion

In this chapter, we proposed an interactive approach based on machine learning for change detection between two landscape images taken from the same location but at different dates. Because of the application dependent significance of a change, as we mentioned in Chapter 3, it is quite difficult to obtain satisfactory results using a fully automatic process. The knowledge of what is a significant change and what is an insignificant change has to be taken into account.

We chose to simply ask the user to designate a few examples of changed and unchanged pixels in the considered image pair. These samples are used to train a classifier so that

it can then label all the remaining pixels in the image pair. Finally, the obtained change maps are post-processed using simple mathematical morphology operators for removing small wrongly labelled regions.

Each pixel is described by a feature vector calculated from the dissimilarity image. The latter is obtained using the SSD distance between the neighborhoods of the two pixels located at the same position in the two images. As a classifier we investigated decision trees and random forests.

In order to evaluate the performance of the proposed method, to compare the two classifiers and to determine the best neighborhood size, we made experiments on the same image pairs we used for testing the automatic method presented in Chapter 3 and we retained the same performance criterions.

Regarding the classifier choice, our experiments show the better performance of random forests but at the price of a longer processing time.

About the neighborhood size for the computation of the SSD dissimilarity image, a small 3×3 size is able to give better results than 5×5 and 7×7 .

Although the interactive method is more constraining for the user, we can conclude that it outperforms significantly the automatic approach we proposed. This superiority can be observed with all the criterions including the dispersion of their values.

Chapter 5

Conclusion and Outlook

The work reported in this dissertation focused on detecting changes from photographs of the same scene taken from the ground at large time intervals. This is a first step towards making available specific image processing tools for the exploitation of photographic landscape observatories.

After a review of existing change detection methods, the main contributions of this work lie in the proposition of two methods. The first one is fully automatic while the second requires help from the user.

For the first method, in order to avoid the detection of insignificant changes between two images, like vegetation variations due to seasonality, we measure the change level of textural features. These features are not calculated for each pixel, but for each superpixel obtained by over-segmenting the two images. Texton histograms completed by the mean gray-level are compared between corresponding superpixels using the Mahalanobis distance. Finally, the change level matrix containing these distance values is binarized into a final change map.

For the second method, we investigated a machine learning approach using supervised classification where training data are given in the form of sample pixels selected by the user. Each pixel is described by a 16-dimensional feature vector derived from a dissimilarity image and the gray-levels of the two images. The feature vectors of the selected sample pixels are used to train a classifier which allows to decide for each remaining pixel whether or not it belongs to a change area. Finally, some small groups of misclassified pixels are corrected by closing and opening the binary change map using a small structuring element.

We evaluated the performance of the proposed methods and their variants through experiments based on 12 image pairs, their corresponding ground truth change maps and six performance criterions. Among the two segmentation methods used for extracting superpixels in the automatic approach, Achanta (SLIC) segmentation gives the best

results. Experiments show that the automatic method outperforms a simple image differencing technique but encounter difficulties for distinguishing significant from insignificant changes.

Performance evaluation of the interactive approach shows that it gives significantly better results than the automatic method. A small neighborhood appears to be sufficient to efficiently measure the local dissimilarity between each pixel of the first image and the pixel located at the same positon in the second image. Among the two classifiers, random forests are more time consuming but produce the best change maps.

This study confirms that the problem of change detection is challenging. We think that it is difficult to avoid any kind of help from the user that is particularly useful for taking into account the semantic of changes. In other words, it seems necessary to learn what change is meaningful or not, because this knowledge highly depends on the application. In a context where it is difficult to model accurately what we want to obtain, machine learning and help from the user allows to solve, at least partially, the problem of change detection.

Of course, some issues remain. First of all, the quality of the results we obtained are far from being perfect. Even if some additional investigations have been conducted but not reported in this document because of their poor improvement ability, further work is necessary to try and get better performance of the change detection method we propose. Below are propositions of some tracks to follow.

- Pixel and superpixel description: Other complementary local descriptors could be used to describe the considered primitives. In addition to textons and co-occurrence matrix-based features, other texture descriptors like those derived from the local binary patterns [PHZA11] could be used. Except for superpixel segmentation, we use only gray levels. So incorporating color descriptors [vdSGS10, DMBC11] in the change detection method we proposed could be another line of investigation.
- Iterative interactive change detection: In the proposed interactive method, for a given image pair, the user gives samples of changed and unchanged classes, the classifier is trained with these data and it classifies all the other pixels of the image pair. As the obtained result can contain errors, we could propose the user to add samples in order to train again the classifier and to improve the result. This sample updating could be repeated so as to become an iterative interactive change detection method.

  In addition, the trained classifier could be used not only on the current image pair, but also on other pairs from the same series or even on other landscape image pairs.
- Semantic change detection: An approach inspired by the "from-to" change detection (post-classification) used in remote sensing applications could be investigated. In addition to the location of changed areas, it would be useful to describe the nature

of the change by extracting semantic information. The method would consist in processing a semantic segmentation of the two images in order to assign meaningful labels to each region and thus to each pixel. But the lack of available training data makes it challenging to semantically segment photographic landscape observatories. Nevertheless, recent work on interactive multi-class segmentation [MCP15] seems to be a promising approach.

Part II

Détection des changements à partir de photographies (version française résumée)

Résumé

Les travaux de cette thèse concernent la détection des changements dans des séries chronologiques de photographies de paysages prises depuis le sol.

Ce contexte de comparaison d'images successives est celui que rencontrent les géographes de l'environnement qui ont recours aux observatoires photographiques du paysage. Ces outils d'analyse et d'aide à la décision sont des bases de données de photographies constituées selon une méthodologie stricte de rephotographie de la même scène, à des pas de temps réguliers. Le nombre de clichés est parfois très important, et l'analyse humaine fastidieuse et relativement imprécise, aussi un outil automatisant la comparaison de photos de paysage deux à deux pour mettre en évidence les changements serait une aide considérable dans l'exploitation des observatoires photographiques du paysage. Bien entendu, les variations dans l'éclairement, la saisonnalité, l'heure du jour, produisent fatalement des clichés entièrement différents à l'échelle du pixel. Notre objectif était donc de concevoir un système robuste face à ces changements mineurs, mais capable de détecter les changements pertinents de l'environnement.

De nombreux travaux autour de la détection des changements ont été effectués pour des images provenant de satellites. Mais l'utilisation d'appareils photographiques numériques classiques depuis le sol pose des problèmes spécifiques comme la limitation du nombre de bandes spectrales et la forte variation de profondeur dans une même image qui induit des apparences différentes des mêmes catégories d'objets en fonction de leurs positions dans la scène.

Dans un premier temps, nous avons exploré la voie de la détection automatique des changements. Nous avons proposé une méthode reposant sur le recalage et la sur-segmentation des images en superpixels. Ces derniers sont ensuite décrits par leur niveau de gris moyen ainsi que par leur texture au travers d'une représentation sous la forme d'histogrammes de textons. La distance de Mahalanobis entre ces descripteurs permet de comparer les superpixels correspondants entre deux images prises à des dates différentes. Nous avons évalué les performances de cette approche sur des images de l'observatoire photographique du paysage constitué lors de la construction de l'autoroute A89. Parmi les méthodes de segmentation utilisées pour produire les superpixels, les expérimentations que nous

avons menées ont mis en évidence le bon comportement de la méthode de segmentation d'Achanta.

La pertinence d'un changement étant fortement liée à l'application visée, nous avons exploré dans un second temps une piste faisant intervenir l'utilisateur. Nous avons proposé une méthode interactive de détection des changements reposant sur une phase d'apprentissage. Afin de détecter les changements entre deux images, l'utilisateur désigne, grâce à un outil de sélection, des échantillons constitués d'ensembles de pixels correspondant à des zones de changement et à des zones d'absence de changement. Chaque couple de pixels correspondants, c'est-à-dire situés au même endroit dans les deux images, est décrit par un vecteur de 16 valeurs principalement calculées à partir de l'image des dissemblances. Cette dernière est obtenue en mesurant, pour chaque couple de pixels correspondants, la dissemblance des niveaux de gris de leurs voisinages. Les échantillons désignés par l'utilisateur permettent de constituer des données d'apprentissage qui sont utilisées pour entraîner un classifieur. Parmi les méthodes de classification évaluées, les résultats expérimentaux montrent que les forêts d'arbres décisionnels donnent les meilleurs résultats sur les séries photographiques que nous avons utilisées.

Mots-clés : analyse d'images, détection des changements, superpixels, texture, apprentissage, classification.

Chapitre 1

Introduction

1.1 Objectifs de la thèse

Les travaux présentés dans ce mémoire traitent de la détection des changements à partir de clichés réalisés – au sol – au moyen d'appareils photographiques numériques. Le contexte d'application de cette étude concerne principalement les observatoires photographiques du paysage. Ces bases de données, qui sont fréquemment utilisées par les géographes de l'environnement, sont constituées de séries d'images obtenues en photographiant la même scène à des pas de temps réguliers et en suivant un protocole rigoureux afin de se placer dans des conditions de prise de vue identiques. La position du trépied, le cadrage de la scène et les caractéristiques techniques de l'appareil (capteur, optique, longueur de focale, ouverture) doivent être les mêmes entre deux prises de vue. Ce protocole a été formalisé par le Ministère de l'Écologie, de l'Énergie, du Développement Durable et de l'Aménagement du Territoire dans un document[1] datant de 2008.

Les paysages sont en constante évolution, du fait de facteurs naturels ou humains. Il est nécessaire de suivre leur évolution afin de les comprendre et d'en rendre compte : c'est l'objectif principal des observatoires photographiques du paysage. Ils sont utilisés pour de l'analyse à court terme (études d'impact pour des projets d'aménagement public) ainsi que pour des analyses à long terme (évolution d'un paysage au cours du temps). Ils constituent ainsi des outils d'aide à la décision.

Notons enfin que la France est signataire de la Convention Européenne du Paysage du Conseil de l'Europe, dont l'objectif est de promouvoir la protection, la gestion et l'aménagement des paysages européens et d'organiser la coopération européenne sur les questions de paysage[2] . Les observatoires sont souvent considérés comme un moyen contribuant à

[1] [Méthode de l'Observatoire photographique du paysage] `http://www.developpement-durable.gouv.fr/IMG/DGALN_methodeOPP.pdf`

[2] [Convention de Florence, 2000] `http://www.coe.int/fr/web/landscape`

respecter les engagements de la Convention.

Avec un grand nombre de sites et des prise de vues fréquentes, le nombre d'images d'un observatoire augmente rapidement et rend l'analyse humaine très pénible et source d'erreurs. C'est pourquoi un outil permettant de comparer automatiquement des couples de photographies de paysages afin de mettre en évidence les changements serait d'un grand secours pour exploiter les observatoires photographiques du paysage.

Comme nous le verrons dans le chapitre 2, de nombreuses études ont été menées sur la détection des changements à partir d'images satellite. Mais les méthodes existantes ne peuvent pas être directement appliquées aux observatoires photographiques du paysage. En effet, l'utilisation d'appareils photographiques numériques classiques depuis le sol pose des problèmes spécifiques comme la limitation du nombre de bandes spectrales (basse résolution spectrale). En outre, les fortes différences de profondeur au sein d'une même image induisent, pour des objets d'une même catégorie, une apparence très différente en fonction de leur position dans la scène.

Par ailleurs, deux difficultés peuvent compliquer la détection des changements :

- le décalage géométrique des images ;
- les changements que l'on qualifiera de pertinents dépendent de l'application considérée : par exemple, l'apparition de végétation sur un sol nu est un changement important pour une étude de reprise de végétation sur un site de construction, alors qu'il ne l'est pas dans le cadre d'une étude d'occupation des sols en agriculture où il n'est dû qu'à la saison.

Heureusement, lors des prises de vues sur lesquelles porte cette étude, la configuration (position du trépied, cadrage de la scène et spécifications de l'appareil) est quasiment la même pour toutes les images d'une même série. Cette contrainte nous permet de disposer d'images presque « alignées ». Une procédure de recalage géométrique des couples d'images est tout de même nécessaire. Cependant, les variations d'éclairage, la saisonnalité ou l'heure de la journée produisent des images complètement différentes à l'échelle du pixel et amènent donc des changements anodins. Notre objectif est de concevoir un système robuste aux changements mineurs mais capable de détecter les changements significatifs.

1.2 Contributions et organisation du mémoire

Dans ce mémoire, nous nous consacrons à la détection des zones de changement dans deux images de la même scène prises par un appareil photographique numérique à des dates différentes. Nous examinons les principales approches existantes et nous proposons deux méthodes pour détecter les changements dans les observatoires photographiques du

paysage. Pour cela, le mémoire est organisé de la manière suivante.

Le chapitre 2 est consacré aux méthodes existantes. Nous détaillons plus particulièrement les méthodes algébriques, les méthodes reposant sur une transformation et celles fondées sur une classification.

Après avoir décrit la procédure de recalage géométrique des images que nous avons mise en place, nous proposons au chapitre 3 une approche automatique de la détection des changements qui s'appuie sur une segmentation des images en « superpixels » et une description de la texture à l'aide des « textons ». Au lieu de considérer les pixels de manière indépendante, nous les regroupons en petites régions. La méthode est constituée de trois étapes principales : la sur-segmentation des images, la description des superpixels grâce à l'extraction des caractéristiques de texture et, enfin, la mise en évidence des changements. Cette dernière est réalisée en mesurant les distances entre les descripteurs des superpixels qui se correspondent entre les deux images. Nous comparons expérimentalement deux méthodes de sur-segmentation permettant d'extraire des superpixels. Enfin, la carte binaire des changements est obtenue par « seuillage » de la matrice contenant, pour chaque pixel son degré de changement.

Bien que l'approche automatique de détection des changements permette d'obtenir des résultats intéressants, nous pensons que leur qualité est insuffisante pour être exploitée dans un observatoire. En effet, des changements non significatifs sont tout de même détectés. Une piste consiste à faire intervenir l'utilisateur pour obtenir de meilleurs résultats. Nous proposons ainsi au chapitre 4 une méthode interactive de détection des changements reposant sur une mesure de dissemblance et une méthode d'apprentissage. Pour décrire chaque pixel, nous utilisons un vecteur d'attributs comprenant des descripteurs de la texture de l'image des dissemblances ainsi que des indices locaux calculés à partir du voisinage du pixel. Grâce à un outil de sélection, l'utilisateur désigne des pixels correspondant à des zones de changement et à des zones d'absence de changement. Ces échantillons sont ensuite utilisés pour entraîner un « classifieur » grâce à une méthode d'apprentissage. Le classifieur est ensuite utilisé pour classer tous les autres pixels du couple d'images.

La carte binaire des changements ainsi obtenue contient des petits groupes de pixels mal classés que nous corrigeons grâce à des opérateurs simples de morphologie mathématique.

Enfin, le chapitre 5 conclut ce mémoire et donne quelques pistes pour poursuivre ce travail.

Chapitre 2

Étude bibliographique de la détection des changements

Une méthode de détection des changements conforme à nos attentes doit être capable de distinguer les changements pertinents des changements non significatifs. Pour une application donnée, il faudra donc concevoir un algorithme prenant en compte les différents types de changement intéressants. MacLeod et Congalton [MC98] ont décrit de la manière suivante les étapes de la détection des changements pour le suivi des ressources naturelles : détecter si un changement s'est produit, identifier sa nature, mesurer son ampleur et estimer la configuration spatiale du changement.

À cause des conséquences de multiples facteurs, il n'est pas facile de déterminer quel algorithme convient pour une application donnée. Une étude des méthodes existantes de détection des changements s'avère donc très utile pour comprendre les principes de fonctionnement de ces méthodes et la manière dont elles peuvent être utilisées pour résoudre différentes catégories de problèmes.

Ce chapitre propose donc un tour d'horizon des méthodes existantes de détection des changements. Ces méthodes sont généralement constituées d'un pré-traitement des images (par exemple un recalage géométrique ou un ajustement radiométrique) suivi de la détection des changements.

Il convient de noter que certaines étapes des méthodes de détection des changements pour la télédétection peuvent s'avérer très utiles dans le contexte des observatoires photographiques du paysage : nous leur accordons une attention particulière.

Chapitre 3

Détection automatique des changements par segmentation en superpixels et description de texture

Dans ce chapitre, nous proposons une approche automatique de détection des changements fondée sur la segmentation des images en superpixels et la description de leur texture. Afin de compenser le manque d'information à l'échelle du pixel, la méthode proposée détecte les changements à partir de groupes de pixels, appelés superpixels, qui constituent des petites régions et qui sont extraits par sur-segmentation des images. L'algorithme proposé est constitué de trois étapes principales. Tout d'abord, un recalage est effectué en supposant que la relation géométrique entre les deux images peut être modélisée par une homographie. Ensuite, les images sont sur-segmentées en superpixels. Les superpixels sont alors décrits par leur niveaux de gris moyens et par des attributs de texture constitués par les histogrammes de textons. Enfin, les degrés de changement sont estimés en comparant les descripteurs des superpixels correspondants. Cette comparaison est effectuée en calculant la distance de Mahalanobis.

Afin d'extraire les superpixels, nous comparons expérimentalement deux algorithmes de segmentation d'images : les méthodes de Shi et d'Achanta. La matrice des changements indiquant le degré de changement de chaque superpixel est donnée par la distance moyenne entre les histogrammes de texton des superpixels correspondants. La carte binaire des changements est alors obtenue en seuillant la matrice des changements.

Comparée à la détection des changements fondée sur la simple différence d'images, la méthode que nous proposons permet d'obtenir des résultats plus précis que nous mettons en évidence dans nos expérimentations.

Chapitre 4

Détection interactive des changements par analyse de l'image des dissemblances et classification supervisée

Cette partie de notre travail est également consacrée à la détection des zones de changement dans deux images de la même scène prises à des dates différentes, mais nous allons prendre en considération un nouvel aspect du problème.

Les images prises par des appareils photographiques numériques contiennent généralement moins d'informations que les images multi-spectrales de télédétection. Par ailleurs, des changements généralement non significatifs pour la majorité des applications, comme ceux qui sont induits par les ombres portées ou les nuages, peuvent mettre en échec les méthodes classiques qui reposent sur la différence d'images.

L'apprentissage automatique est une piste prometteuse mais la faible quantité de données d'apprentissage disponibles dans un observatoire photographique du paysage disqualifie de nombreuses méthodes. Aussi, allons-nous explorer dans ce chapitre la voie de l'apprentissage interactif.

Afin de décrire chaque pixel, nous construisons un vecteur de 16 attributs calculés à partir d'un petit voisinage. La majorité des attributs sont calculés à partir de l'image des dissemblances qui exprime la distance entre les voisinages de chaque pixel dans la première et dans la seconde images. Le classifieur est entraîné à partir des échantillons désignés par l'utilisateur grâce à un outil de sélection. Nous avons comparé expérimentalement deux types de classifieurs : les arbres de décision et les forêts d'arbres décisionnels. Le classifieur est ensuite utilisé pour étiqueter tous les autres pixels du couple. La carte binaire des changements ainsi obtenue subit enfin un post-traitement avec des opérateurs simples de

morphologie mathématique afin de supprimer des petits groupes de pixels mal classés, c'est-à-dire qui sont considérés à tort comme des éléments d'une zone de changement (faux positifs) ou d'absence de changement (faux négatifs).

Les résultats expérimentaux, que nous avons obtenus en suivant le même protocole que pour l'évaluation des performances de la méthode automatique présentée au chapitre précédent, montrent l'intérêt de l'approche proposée.

Chapitre 5

Conclusion et perspectives

Dans cette thèse, nous nous sommes consacrés à la détection des changements à partir de photographies de la même scène prises depuis le sol à de grands intervalles de temps. Les observatoires photographiques du paysage ont constitué le contexte d'application de ce travail. Nous avons exploré les approches existantes et nous avons proposé deux nouvelles méthodes de détection des changements.

Après avoir proposé une procédure de recalage géométrique des couples d'images, nous avons décrit une méthode automatique pour identifier les zones de changement entre deux images. Afin de compenser le manque d'informations des images prises par des appareils photographiques, nous avons remplacé le traitement des pixels indépendamment les uns des autres par des groupes de pixels. Nous avons proposé de caractériser ces groupes, appelés superpixels, par un descripteur constitué de leur niveau de gris moyen et l'histogramme de textons qui constitue l'un des moyens de décrire leur texture. Nous avons évalué expérimentalement les performances de la méthode en utilisant deux méthodes de segmentation en superpixels. La méthode de segmentation d'Achanta (SLIC) a donné les résultats les plus satisfaisants sur les images que nous avons utilisées pour cette évaluation.

Nous avons également proposé une approche interactive de la détection des changements reposant sur l'image des dissemblances et une classification supervisée. Nous avons expérimenté deux méthodes de classification associées à la mesure de dissemblance SSD. Cette approche interactive a permis de détecter la plupart des zones de changement en étant robuste face aux conséquences de petites erreurs de recalage ainsi qu'aux effets des ombres. Cependant, l'extraction précise de changements fins peut parfois échouer. Les expérimentations que nous avons menées ont mis en évidence les bons résultats obtenus avec les forêts d'arbres décisionnels associées à un voisinage de taille 3×3 qui ont ainsi donné les meilleurs résultats.

Les perspectives de ce travail se situent à la fois dans l'amélioration des méthodes proposées et dans l'exploration d'approches alternatives.

- Description des pixels et des superpixels :
  Des descripteurs locaux complémentaires pourraient être utilisés pour caractériser les primitives. En plus des textons et des matrices de co-occurrence, d'autres descripteurs de texture comme ceux de la famille des motifs binaires locaux [PHZA11] pourraient être envisagés. Par ailleurs, sauf pour la segmentation des images en superpixels, nous n'avons utilisé que les niveaux de gris. La prise en compte de descripteurs de couleur [vdSGS10, DMBC11] dans la méthode de détection des changements constitue une autre piste à explorer.
- Détection du changement interactive et itérative :
  Dans la méthode interactive que nous avons proposée, pour un couple d'images, l'utilisateur sélectionne des échantillons des deux classes « changement » et « absence de changement », le classifieur est ensuite entraîné avec ces données et il classe tous les autres pixels du couple. Comme le résultat obtenu peut contenir des erreurs, nous pourrions proposer à l'utilisateur d'ajouter des échantillons en fonction des erreurs qu'il observe afin d'entraîner de nouveau le classifieur et d'améliorer la qualité des résultats. Cette mise à jour des échantillons pourrait être répétée afin de constituer une méthode interactive itérative de détection des changements. En outre, le classifieur entraîné pourrait être utilisé non seulement sur le couple d'images, mais aussi sur d'autres couples de la même série ou même sur d'autres couples d'images de paysage.
- Détection sémantique des changements :
  Une approche inspirée de la détection des changement « de-vers » (post-classification) utilisée dans les applications de télédétection pourrait être envisagée. En plus de la localisation des zones de changement, il pourrait être utile de décrire la nature des changements en extrayant des informations sémantiques. La méthode consisterait à effectuer une segmentation sémantique des deux images afin d'attribuer à chaque région, et donc à chaque pixel, une étiquette porteuse de sens. Mais le manque de données d'apprentissage disponibles complique la segmentation sémantique des images des observatoires photographiques du paysage. Néanmoins, des travaux récents autour de la segmentation interactive multi-classes [MCP15] semblent constituer une piste prometteuse.

Bibliography

[AGSD12] J. Aguirre-Gutiérrez, A. C. Seijmonsbergen, and J. F. Duivenvoorden. Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. *Applied Geography*, 34:29–37, 2012.

[Ahl08] O. Ahlqvist. Extending post-classification change detection using semantic similarity metrics to overcome class heterogeneity: A study of 1992 and 2001 US national land cover database changes. *Remote Sensing of Environment*, 112(3):1226–1241, 2008.

[ASS+10] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC superpixels. Technical report, EPFL, 2010.

[ASS+12] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012.

[BA97] L. A. Breslow and D. W. Aha. Simplifying decision trees: A survey. *The Knowledge Engineering Review*, 12(01):1–40, 1997.

[BB07] F. Bovolo and L. Bruzzone. A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment. *IEEE Transactions on Geoscience and Remote Sensing*, 45(6):1658–1670, 2007.

[BCA13] A. Belghith, C. Collet, and J. P. Armspach. Change detection based on a support vector data description that treats dependency. *Pattern Recognition Letters*, 34(3):275–282, 2013.

[BCM80] G. F. Byrne, P. F. Crapper, and K. K. Mayo. Monitoring land-cover change by principal component analysis of multitemporal Landsat data. *Remote Sensing of Environment*, 10(3):175–184, 1980.

[BFSO84] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and regression trees*. CRC press, 1984.

[BOC+13] S. Boukir, C. Orny, N. Chehata, D. Guyon, and J. P. Wigneron. Détection de changements structurels sur des images satellite haute résolution: Application en milieu forestier. *TS. Traitement du signal*, 30(6):401–429, 2013.

[BP00a] L. Bruzzone and D. F. Prieto. An adaptive parcel-based technique for unsupervised change detection. *International Journal of Remote Sensing*, 21(4):817–822, 2000.

[BP00b] L. Bruzzone and D. F. Prieto. Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 38(3):1171–1182, 2000.

[BP02] L. Bruzzone and D. F. Prieto. A partially unsupervised cascade classifier for the analysis of multitemporal remote-sensing images. *Pattern Recognition Letters*, 23(9):1063–1071, 2002.

[Bre01] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[Bro92] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[BS97] L. Bruzzone and S. B. Serpico. An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):858–867, 1997.

[CB96] P. R. Coppin and M. E. Bauer. Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing Reviews*, 13(3-4):207–234, 1996.

[CC11] S. Chambon and A. Crouzil. Similarity measures for image matching despite occlusions in stereo vision. *Pattern Recognition*, 44(9):2063–2075, 2011.

[CDP07] L. Castellana, A. D'Addabbo, and G. Pasquariello. A composed supervised/unsupervised approach to improve change detection from remote sensing. *Pattern Recognition Letters*, 28(4):405–413, 2007.

[Cel09] T. Celik. Unsupervised change detection in satellite images using principal component analysis and k-means clustering. *Geoscience and Remote Sensing Letters*, 6(4):772–776, 2009.

[CFMC01] L. M. T. de Carvalho, L. M. G. Fonseca, F. Murtagh, and J. G. P. W. Clevers. Digital change detection with the aid of multiresolution wavelet analysis. *International Journal of Remote Sensing*, 22(18):3871–3876, 2001.

[CGH+03] J. Chen, P. Gong, C. He, R. Pu, and P. Shi. Land-use/land-cover change detection using improved change-vector analysis. *Photogrammetric Engineering and Remote Sensing*, 69(4):369–379, 2003.

[CJK89] P. S. Chavez Jr and A. Y. Kwarteng. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogrammetric Engineering and Remote Sensing*, 55(3):339–348, 1989.

[CLG89] V. Caselles and M. J. Lopez Garcia. An alternative simple approach to estimate atmospheric correction in multitemporal studies. *International Journal of Remote Sensing*, 10(6):1127–1134, 1989.

[CNS04] M. J. Canty, A. A. Nielsen, and M. Schmidt. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sensing of Environment*, 91(3):441–451, 2004.

[COB+14] N. Chehata, C. Orny, S. Boukir, D. Guyon, and J. P. Wigneron. Object-based change detection in wind storm-damaged forest using high-resolution multispectral images. *International Journal of Remote Sensing*, 35(13):4758–4777, 2014.

[COBG11] N. Chehata, C. Orny, S. Boukir, and D. Guyon. Object-based forest change detection using high resolution satellite images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII*, pages 5–7, 2011.

[DK98] X. Dai and S. Khorram. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):1566–1577, 1998.

[DLR77] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39:1–38, 1977.

[DMBC11] G. Demarcq, L. Mascarilla, M. Berthier, and P. Courtellemont. The color monogenic signal: Application to color edge detection and color optical flow. *Journal of Mathematical Imaging and Vision*, 40(3):269–284, 2011.

[DTC02] Y. Du, P. M. Teillet, and J. Cihlar. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote sensing of Environment*, 82(1):123–134, 2002.

[DTJ+10] M. Ding, Z. Tian, Z. Jin, M. Xu, and C. Cao. Registration using robust kernel principal component for object-based change detection. *Geoscience and Remote Sensing Letters*, 7(4):761–765, 2010.

[EG06] P. Enjelvin and C. Guy. Observatoire photographique des territoires du Massif Central. `http://poptmc.free.fr`, 2006.

[EM03] F. Eugenio and F. Marqués. Automatic satellite image georeferencing using a contour-matching approach. *IEEE Transactions on Geoscience and Remote Sensing*, 41(12):2869–2880, 2003.

[ES93] L. Eklundh and A. Singh. A comparative analysis of standardised and unstandardised principal components analysis in remote sensing. *International Journal of Remote Sensing*, 14:1359–1370, 1993.

[FL87] T. Fung and E. LeDrew. Application of principal components analysis to change detection. *Photogrammetric Engineering and Remote Sensing*, 53(12):1649–1658, 1987.

[FPS08] F. D. Frate, F. Pacifici, and D. Solimini. Monitoring urban land cover in rome, italy, and its changes by single-polarization multitemporal SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(2):87–97, 2008.

[GLM92] P. Gong, E. F. LeDrew, and J. R. Miller. Registration-noise reduction in difference images for change detection. *International Journal of Remote Sensing*, 13(4):773–779, 1992.

[GSJC14] M. Gong, L. Su, M. Jia, and W. Chen. Fuzzy clustering with a modified MRF energy function for change detection in synthetic aperture radar images. *IEEE Transactions on Fuzzy Systems*, 22(1):98–109, 2014.

[GSP86] A. Goshtasby, G. C. Stockman, and C. V. Page. A region-based approach to digital image registration with subpixel accuracy. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24(3):390–399, 1986.

[HHSMA98] T. Hame, I. Heiler, and J. San Miguel-Ayanz. An unsupervised change detection and recognition system for forestry. *International Journal of Remote Sensing*, 19(6):1079–1099, 1998.

[HS01] D. J. Hayes and S. A. Sader. Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series. *Photogrammetric Engineering and Remote Sensing*, 67(9):1067–1075, 2001.

[HSD73] R. M. Haralick, K. Shanmugam, and I. H. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6):610–621, 1973.

[HSNG91] F. G. Hall, D. E. Strebel, J. E. Nickeson, and S. J. Goetz. Radiometric rectification: toward a common radiometric response among multidate, multisensor images. *Remote Sensing of Environment*, 35(1):11–27, 1991.

[HWS+11] C. He, A. Wei, P. Shi, Q. Zhang, and Y. Zhao. Detecting land-use/land-cover change in rural-urban fringe areas using extended change-vector analysis. *International Journal of Applied Earth Observation and Geoinformation*, 13(4):572–585, 2011.

[IJ05] J. Im and J. R. Jensen. A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sensing of Environment*, 99(3):326–340, 2005.

[IJT08] J. Im, J. R. Jensen, and J. A. Tullis. Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing*, 29(2):399–423, 2008.

[IM07] J. Inglada and G. Mercier. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5):1432–1445, 2007.

[Jen86] J. R. Jensen. Introductory digital image processing: a remote sensing perspective. *Englewood Cliffs, NJ: Prentice-Hall*, 1986.

[Jen05] J. R. Jensen. *Introductory digital image processing: A remote sensing perspective*. Pearson College Division, 2005.

[JK10] R. D. Johnson and E. S. Kasischke. Change vector analysis: A technique for the multispectral monitoring of land cover and condition. *International Journal of Remote Sensing*, pages 411–426, 2010.

[Jul81] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802):91–97, 1981.

[LLJ02] X. Liu and R. G. Lathrop Jr. Urban change detection based on an artificial neural network. *International Journal of Remote Sensing*, 23(12):2513–2518, 2002.

[LMBM04] D. Lu, P. Mausel, E. Brondizio, and E. Moran. Change detection techniques. *International Journal of Remote Sensing*, 25(12):2365–2401, 2004.

[LMCC02] J. Le Moigne, W. J. Campbell, and R. F. Cromp. An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8):1849–1864, 2002.

[LMM95] H. Li, B. S. Manjunath, and S. K. Mitra. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4(3):320–334, 1995.

[LYLE98] J. G. Lyon, D. Yuan, R. S. Lunetta, and C. D. Elvidge. A change detection experiment using vegetation indices. *Photogrammetric Engineering and Remote Sensing*, 64(2):143–150, 1998.

[Mal80] W. A. Malila. Change vector analysis: An approach for detecting forest changes with landsat. In *LARS Symposia*, pages 325–335, 1980.

[MBLS01] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 43(1):7–27, 2001.

[MC98] R. D. Macleod and R. G. Congalton. A quantitative comparison of change-detection algorithms for monitoring Eelgrass from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 64(3):207–216, 1998.

[MCP15] B. Mathieu, A. Crouzil, and J-B. Puel. Interactive multiclass segmentation using superpixel classification. *Computing Research Repository*, abs/1510.03199, 2015.

[MH94] D. M. Muchoney and B. N. Haack. Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, 60(10):1243–1251, 1994.

[Mit97] T. M. Mitchell. Decision tree learning, in: T. Mitchell (Ed.), Machine Learning. *McGraw Hill*, pages 52–78, 1997.

[NC01] I. Niemeyer and M. J. Canty. Object-oriented post-classification of change images. In *Proceedings of the SPIE's International Symposium on Remote Sensing*, volume 4545, pages 100–108, 2001.

[NC03] I. Niemeyer and M. J. Canty. Pixel-based and object-oriented change detection analysis using high-resolution imagery. In *Proceedings of 25th Symposium on Safeguards and Nuclear Material Management*, pages 2133–2136, 2003.

[PHZA11] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen. *Computer Vision Using Local Binary Patterns*. Computational imaging and vision. Springer Verlag, 2011.

[PM01] M. Pal and P. M. Mather. Decision tree based classification of remotely sensed data. In *Proceedings of 22nd Asian Conference on Remote Sensing*, volume 5, page 9, 2001.

[PSL01] C. Petit, T. Scudder, and E. Lambin. Quantifying processes of land-cover change by remote sensing: resettlement and rapid land-cover changes in south-eastern Zambia. *International Journal of Remote Sensing*, 22(17):3435–3456, 2001.

[PT98] P. Perner and S. Trautzsch. Multi-interval discretization methods for decision tree learning. In *Advances in Pattern Recognition*, pages 475–482. Springer, 1998.

[Qui86] J. R. Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.

[Qui93] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993.

[RAAKR05] R. J. Radke, S Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing*, 14(3):294–307, 2005.

[Ros98] P. L. Rosin. Thresholding for change detection. In *Proceedings of Sixth International Conference on Computer Vision*, pages 274–279, 1998.

[Rou84] P. J. Rousseeuw. Least median of squares regression. *Journal of the American statistical association*, 79(388):871–880, 1984.

[SA00] P. C. Smits and A. Annoni. Toward specification-driven change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 38(3):1484–1488, 2000.

[SCM90] D. A. Stow, D. Collins, and D. McKinsey. Land use change detection based on multi-date imagery from different satellite sensor systems. *Geocarto International*, 5(3):3–12, 1990.

[Ser82] J. Serra. *Image analysis and mathematical morphology, v. 1.* Academic press, 1982.

[Sin89] A. Singh. Digital change detection techniques using remotely sensed data. *International Journal of Remote Sensing*, 10(6):989–1003, 1989.

[SL91] S. R Safavian and D. Landgrebe. A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3):660–674, 1991.

[SM00] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[Soh99] T. L. Sohl. Change analysis in the United Arab Emirates: an investigation of techniques. *Photogrammetric Engineering and Remote Sensing*, 65(4):475–484, 1999.

[SSS08] Y. Sheng, C. Shah, and L. C. Smith. Automated image registration for hydrologic change detection in the lake-rich Arctic. *Geoscience and Remote Sensing Letters*, 5(3):414–418, 2008.

[ST94] J. Shi and C. Tomasi. Good features to track. In *1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994. Proceedings*, pages 593–600. IEEE, 1994.

[ST99] L. K. Soh and C. Tsatsoulis. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2):780–795, 1999.

[Sto99] D. A. Stow. Reducing the effects of misregistration on pixel-level change detection. *International Journal of Remote Sensing*, 20(12):2477–2483, 1999.

[TJGM92] J. R. G. Townshend, C. O. Justice, C. Gurney, and J. McManus. The impact of misregistration on change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 30(5):1054–1060, 1992.

[VB00] D. L. Verbyla and S. H. Boles. Bias in land cover change estimates due to misregistration. *International Journal of Remote Sensing*, 21(18):3553–3560, 2000.

[vdSGS10] K. van de Sande, T. Gevers, and C. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 2010.

[VTB+13] M. Volpi, D. Tuia, F. Bovolo, M. Kanevski, and L. Bruzzone. Supervised change detection in VHR images using contextual information and support vector machines. *International Journal of Applied Earth Observation and Geoinformation*, 20:77–85, 2013.

[VTCVK12] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski. Unsupervised change detection with kernels. *Geoscience and Remote Sensing Letters*, 9(6):1026–1030, 2012.

[Wal04] V. Walter. Object-based classification of remote sensing data for change detection. *International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing*, 58(3):225–238, 2004.

[Web01] K. T. Weber. A method to incorporate phenology into land cover change analysis. *Journal of Range Management*, 54(2):A1–A7, 2001.

[WMPLC01] C. E. Woodcock, S. A. Macomber, M. Pax-Lenney, and W. B. Cohen. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment*, 78(1):194–203, 2001.

[YB15] O. Yousif and Y. Ban. Object-based urban change detection using high resolution SAR images. In *Proceedings on Urban Remote Sensing Event*, pages 1–4, 2015.

[YL00] X. Yang and C. P. Lo. Relative radiometric normalization performance for change detection from multi-date satellite images. *Photogrammetric Engineering and Remote Sensing*, 66(8):967–980, 2000.

[ZF03] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.