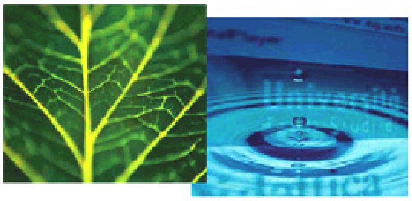
PhD Dissertation



International Doctorate School in Information and Communication Technologies

DISI - University of Trento

ADVANCED TECHNIQUES FOR AUTOMATIC CHANGE DETECTION IN MULTITEMPORAL HYPERSPECTRAL IMAGES

Sicong Liu

Advisor:

Prof. Lorenzo Bruzzone

Università degli Studi di Trento

Co-Advisor:

Prof. Peijun Du

Nanjing University

Dr. Francesca Bovolo

Fondazione Bruno Kessler

February 2015

If I have seen further, it is by standing on the shoulders of giants.

-Isaac Newton

Dedicated to my mother.

Abstract

The increasing availability of the new generation remote sensing satellite hyperspectral images provides an important data source for Earth Observation (EO). Hyperspectral images are characterized by a very detailed spectral sampling (i.e., very high spectral resolution) over a wide spectral wavelength range. This important property makes it possible the monitoring of the land-cover dynamic and environmental evolution at a fine spectral scale. This also allows one to potentially detect subtle spectral variations associated with the land-cover transitions that are usually not detectable in the traditional multispectral images due to their poor spectral signature representation (i.e., generally sufficient for representing only the major changes). To fully utilize the available multitemporal hyperspectral images and their rich information content, it is necessary to develop advanced techniques for robust change detection (CD) in multitemporal hyperspectral images, thus to automatically discover and identify the interesting and valuable change information. This is the main goal of this thesis.

In the literature, most of the CD approaches were designed for multispectral images. The effectiveness of these approaches, to the complex CD problems is reduced, when dealing with the hyperspectral images. Accordingly, the research activities carried out during this PhD study and presented in this thesis are devoted to the development of effective methods for multiple change detection in multitemporal hyperspectral images. These methods consider the intrinsic properties of the hyperspectral data and overcome the drawbacks of the existing CD techniques. In particular, the following specific novel contributions are introduced in this thesis:

1) *A theoretical and empirical analysis of the multiple-change detection problem in multitemporal hyperspectral images. Definition and discussion of concepts as the changes and of the change endmembers, the hierarchical change structure and the multitemporal spectral mixture is given.*
2) *A novel semi-automatic sequential technique for iteratively discovering, visualizing, and detecting multiple changes. Reliable change variables are adaptively generated for the representation of each specific considered change. Thus multiple changes are discovered and discriminated according to an iterative re-projection of the spectral change vectors into new compressed change representation domains. Moreover, a simple yet effective tool is developed allowing user to have an interaction within the CD procedure.*
3) *A novel partially-unsupervised hierarchical clustering technique for the separation and identification of multiple changes. By considering spectral variations at different processing levels, multiple change information is adaptively modelled and clustered according to spectral homogeneity. A manual initialization is used to drive the whole hierarchical clustering procedure;*
4) *A novel automatic multitemporal spectral unmixing approach to detect multiple changes in hyperspectral images. A multitemporal spectral mixture model is proposed to analyse the spectral variations at sub-pixel level, thus investigating in details the spectral composition of change and no-*

change endmembers within a pixel. A patch-scheme is used in the endmembers extraction and unmixing, which better considers endmember variability.

Comprehensive qualitative and quantitative experimental results obtained on both simulated and real multitemporal hyperspectral images confirm the effectiveness of the proposed techniques.

Keywords

Change detection (CD), change visualization, change representation, change vector analysis, spectral unmixing, hyperspectral (HS) images, multiple changes, multitemporal analysis, remote sensing, hierarchical analysis.

Contents

List of Figures

List of Tables

List of Abbreviations

EO	Earth Observation
RS	Remote Sensing
SAR	Synthetic Aperture Radar
LiDAR	Light Detection And Ranging
HS	Hyperspectral
MS	Multispectral
CD	Change Detection
CD-MS	Change Detection in Multispectral images
CD-HS	Change Detection in Hyperspectral images
S^2CVA	Sequential Spectral Change Vector Analysis
MSU	Multitemporal Spectral Unmixing
PCC	Post-Classification Comparison
DMC	Direct Multi-date Classification
SVM	Support Vector Machine
PCA	Principal Component Analysis
FCM	Fuzzy C-Means
GKC	Gustafson-Kessel Clustering
GA	Genetic Algorithm
SA	Simulated Annealing
MRF	Markov Random Fields
GS	Gramm-Schmidt
KT	Tasselled-cap Transformation
MAD	Multivariate Alteration Detection
CCA	Canonical Correlation Analysis
IR-MAD	Iterative Reweighted MAD
CVA	Change Vector Analysis
C^2VA	Compressed Change Vector Analysis
MAPs	Morphological Attribute Profiles
SFA	Slow Feature Analysis
CE	Covariance-Equalization
QCE	Class-conditional CE
WDS	Wavelength Dependent Segmentation
MAD	Multivariate Alteration Detection
MAF	Maximum Autocorrelation Factor
MNF	Minimum Noise Fraction
T-PCA	Temporal-Principal Component Analysis
ICA	Independent Component Analysis
UFD	Uniform Feature Design
stICA	Spatio-temporal ICA
SAM	Spectral Angle Measure
SID	Spectral Information Divergence
SCM	Spectral Correlation Measure
HOOI	Higher Order Orthogonal Iteration
SNR	Signal-to-Noise Ratio
SCV	Spectral Change Vector
MKT	Multi-date Kauth-Thomas
MGS	Multi-date Graham-Schimidt
MT-EMs	Multitemporal Endmembers
LMM	Linear Mixture Model
SAD	Spectral Angle Distance
ASCVR	Adaptive Spectral Change Vector Representation

SVD	Singular Value Decomposition
USGS	U.S. Geological Survey
C^2VA_T	Automatic Thresholding in the C^2VA feature space
C^2VA_M	Manual (interactive) change identification in the C^2VA feature space
k-means_SCVs	*k*-means clustering on the whole changed SCVs
S^2CVA_M	S^2CVA approach using Manual (interactive) change identification at each level of the representation
OA	Overall Accuracy
Kappa	Kappa Coefficient
EM	Expectation Maximization
HSCVA	Hierarchical Spectral Change Vector Analysis
PCs	Principal Components
BIC	Bayesian Information Criterion
AIC	Akaike Information Criterion
MDL	Minimum Description Length
MU	Multitemporal Unmixing
HFC	Harsanyi-Farrand-Chang
NWHFC	Noise Whitened Version of HFC
HySIME	Hyperspectral Signal Identification by Minimum Error
ELM	Eigenvalue Likelihood Maximization
VCA	Vertex Component Analysis
ANC	Abundance Nonnegative Constraint

List of Notations

I	Length of the image
J	Width of the image
B	Number of the spectral channels (bands) of the considered images
p	Number of targets (classes)
$H_f(\cdot)$	Hidden function
V_R	Distinct terms in documents
L_D	Length of documents
W, η	Free parameters in the formulation of Heaps' law
$\boldsymbol{X}_1$, $\boldsymbol{X}_2$	Hyperspectral images acquired at time 1 and 2, respectively
$\boldsymbol{X}_D$	Spectral difference image
$x_1(i, j)$	Pixel with spatial position (i, j) in $\boldsymbol{X}_1$
$x_2(i, j)$	Pixel with spatial position (i, j) in $\boldsymbol{X}_2$
Ω	All classes present in the considered images
Ω_c	Set of multiple change classes
ω_n	No-change class
ω_{c_k}	k-th change classes in Ω_c
K	Number of change classes in Ω_c
Ω_e	Set of change endmembers
e_k	k-th change endmember associated to change class ω_{c_k}
$\boldsymbol{X}_S$	Multitemporal stacked image
$\boldsymbol{E}_S$	Matrix of the multitemporal-endmember (MT-EM) set in $\boldsymbol{X}_S$
$\boldsymbol{A}_S$	Abundance matrix of $\boldsymbol{E}_S$ in the linear mixture model
$\boldsymbol{N}_S$	Noise matrix in the linear mixture model
$\boldsymbol{E}_1$, $\boldsymbol{E}_2$	Matrix of endmember set in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively
$\boldsymbol{A}_1$, $\boldsymbol{A}_2$	Abundance matrix of $\boldsymbol{E}_1$, $\boldsymbol{E}_2$ in a linear mixture model, respectively
$\boldsymbol{N}_1$, $\boldsymbol{N}_2$	Noise matrix in the linear mixture model of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively
ρ	Change magnitude
α	Change direction
$\boldsymbol{D}$	2-Dimensional polar coordinate domain
ρ_{max}	Maximum value of ρ
T_ρ	Threshold value on the change magnitude ρ
SC_n	Region associated to the unchanged SCVs in the C^2VA representation
SA_c	Region associated to the changed SCVs in the C^2VA representation
$T_{\alpha,k}$	k-th threshold value on the change direction α
R	Reference vector
L_h	h-th level of S^2CVA hierarchy
J_h	Maximum number of change clusters at level L_h
H	Total number of levels in the S^2CVA hierarchy
$P_{h,j}$	j-th change cluster observed in ASCVR at the level L_h
$\boldsymbol{R}_{h,j}$	Reference vector of ASCVR for cluster $P_{h,j}$
$\boldsymbol{X}_{h,j}$	SCVs associated to cluster $P_{h,j}$
$\boldsymbol{\Gamma}_{h,j}$	Covariance matrix of $\boldsymbol{x}_{h,j}$
$\boldsymbol{E}[\boldsymbol{x}_{h,j}]$	Expectation of $\boldsymbol{x}_{h,j}$
$\boldsymbol{W}_{h,j}$	Diagonal matrix with eigenvalues
$\boldsymbol{V}_{h,j}$	Matrix of eigenvectors
$\lambda_{h,j}^b$	b-th eigenvalue sorted in descending order in $D_{h,j}$
$\boldsymbol{D}_{h,j}$	2-Dimensioanl representation domain for $P_{h,j}$
$\rho_{h,j}$	Constructed change magnitude variable for $P_{h,j}$

$\alpha_{h,j}$	Constructed change direction variable for $P_{h,j}$
e_n	No-change endmember
Ω_u	Set of uncertain pixels
δ	A margin value on the threshold T_ρ
$h(\rho)$	Histogram of magnitude ρ
$\rho(i,j)$	Change magnitude value for pixel $x(i,j)$
L_d	d-th level of the HSCVA tree structure
D	Depth of the HSCVA tree
$\boldsymbol{S}_{\Omega_c}$	Reference spectrum calculated by averaging all SCVs in Ω_c
$\vartheta(\boldsymbol{x}(i,j),\boldsymbol{S}_{\Omega_c})$	Spectral angle distance between $\boldsymbol{x}(i,j)$ and $\boldsymbol{S}_{\Omega_c}$
$\sigma_{\vartheta_{\Omega_c}}$	Standard deviation value of $\vartheta(\boldsymbol{x}(i,j),\boldsymbol{S}_{\Omega_c})$
T_σ	Threshold value on $\sigma_{\vartheta_{\Omega_c}}$
$\boldsymbol{Q}(i,j)$	Selected principal components with spatial position (i,j)
M	Number of the selected PCs
H	Range of number k
k_0	Initial number of k (i.e., lower bound of H)
t	A constant value to control the upper bound of H
$\boldsymbol{Q}_k$	Pixel data of PCs belong to class ω_{c_k}
$f(\cdot)$	probability density function
$\boldsymbol{\Theta}_k$	M-dimensional normal distribution
$\boldsymbol{\mu}_k$	M-dimensional mean vector
$\boldsymbol{\Gamma}_k$	$M \times M$ dimensional covariance matrix
n_k	Number of pixel in $\boldsymbol{Q}_k$
L_f	Likelihood function
$\hat{\boldsymbol{\Theta}}_k$, $\hat{\boldsymbol{\mu}}_k$, $\hat{\boldsymbol{\Gamma}}_k$	Maximum likelihood estimate of $\boldsymbol{\Theta}_k$, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Gamma}_k$
L'_f	Likelihood function of the joint probability density function
k'	Optimal number of major changes
k_t	Temporary number of major changes
$\varphi(\boldsymbol{Q}(i,j))$	Compressed change direction for $\boldsymbol{Q}(i,j)$
$\boldsymbol{S}_{e_\varepsilon}$	Reference spectrum of endmember $e_\varepsilon \in \{\Omega_e, e_n\}$
$B_{\alpha,\beta}$	Bhattacharyya distance between e_α and e_β
$\boldsymbol{\mu}_\alpha$	Mean vector of e_α
$\boldsymbol{\Gamma}_\alpha$	Covariance matrix of e_α
κ	Kappa coefficient
P_1	Number of endmembers in $\boldsymbol{X}_1$
P_2	Number of endmembers in $\boldsymbol{X}_2$
P_s	Number of MT-EMs in $\boldsymbol{X}_S$
$\boldsymbol{X}_{S,z}$	z-th patch of $\boldsymbol{X}_S$
Z	Defined number of patches
$P_{s,z}$	Number of MT-EMs in patch $\boldsymbol{X}_{S,z}$
$\boldsymbol{E}_{S,z}$	Identified set of MT-EMs in patch $\boldsymbol{X}_{S,z}$
$\boldsymbol{U}$	Endmember pool
$\boldsymbol{e}_p$	Spectral signature of the p-th (p=1,…, P_s) endmember in $\boldsymbol{X}_S$
a_p	Fractional abundance of e_p
$\boldsymbol{U}_c$	Endmember pool of change classes
$\boldsymbol{U}_n$	Endmember pool of no-change classes
Ω_n	Set of no-change classes
$P_{s,c}$	Number of endmembers in $\boldsymbol{U}_c$
$P_{s,n}$	Number of endmembers in $\boldsymbol{U}_n$
$\boldsymbol{r}$	Probability vector of a given MT-EM $\boldsymbol{e}_\alpha$

$\boldsymbol{m}$	Probability vector of a given MT-EM $\boldsymbol{e}_\beta$
θ	SID-SAM combined spectral measure
K'	Total number of classes of MT-EMs in $\boldsymbol{X}_S$
ω_ε	A given class in $\Omega = \left\{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}, \omega_n\right\}$
$\boldsymbol{A}_{S,\varepsilon}$	Final abundance map of the class ω_ε
$\boldsymbol{A}_{S,p}$	Abundance map of a given MT-EM $\boldsymbol{e}_p$, $\boldsymbol{e}_p \in \omega_\varepsilon$
$a_{\omega_\varepsilon}(i,j)$	Abundance value of class ω_ε in pixel $x_s(i,j)$
Δ	Euclidean distance
ϑ	Spectral angle distance
a, b	Two constant values
$\boldsymbol{U}^s_{h,j}$, $\boldsymbol{V}^s_{h,j}$	Unitary matrices represent sets of 'left' and 'right' orthonormal bases in SVD
$\left(\boldsymbol{V}^s_{h,j}\right)^*$	Conjugate transpose of the unitary matrix $\boldsymbol{V}^s_{h,j}$
$\boldsymbol{D}^s_{h,j}$	Diagonal matrix in SVD

Chapter 1

Introduction

In this chapter the basic concepts of remote sensing systems, change detection techniques, and hyperspectral sensors are briefly overviewed. The considered challenging multiple change detection problem in multitemporal hyperspectral images is introduced by comparing with the same problem on the traditional multispectral images. Then, the main motivations, objectives and the novel contributions of this thesis are presented. Finally, the whole structure and organization of the thesis are described.

1.1 Introduction to Hyperspectral Remote Sensing Systems

Remote Sensing (RS) is a technique that can continuously observe the Earth surface using a sensor mounted on an aircraft or a spacecraft platform [1]. RS sensors are capable to take measure on the real objects or on the environmental phenomena without having a physical contact with them. This has already been widely used in different application domains (e.g., forest, agriculture, urban areas, ocean, natural disaster, etc.). Depending on the way that the signal is generated, RS can be divided into two main groups: active remote sensing, where the signal is emitted from the sensor (e.g., the Synthetic Aperture Radar (SAR) and Light Detection And Ranging (LiDAR) systems); and the passive remote sensing, where the portion of sunlight radiation reflected from the objects is measured by passive sensors (e.g., optical sensors like photography, infrared, charge-coupled devices and radiometers). In this thesis, the focus is on the study of techniques for addressing the complex and challenging CD procedure in multitemporal hyperspectral (HS) images.

The past few years have witnessed a huge increase in studying hyperspectral images and their applications in different fields. The hyperspectral sensors on board of air crafts (e.g., HYDICE[1] and AVIRIS[2]) or spaceborne-satellites (e.g., Hyperion[3], CHRIS[4], HJ-1A/B[5], IASI[6]), and the ones in the launch schedule (e.g., EnMAP[7], PRISMA[8], HISUI[9], HyspIRI[10]) are providing more and more available hyperspectral data with an increased data quality. In TABLE 1-1, eight hyperspectral instruments together with their spatial and spectral parameters [2] are illustrated, where the EnMAP, PRISMA and HyspIRI are not yet opera-

[1] http://rsd-www.nrl.navy.mil/hydice
[2] http://aviris.jpl.nasa.gov
[3] http://eo1.usgs.gov/sensors/hyperion
[4] https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/proba/instruments/chris
[5] http://www.cresda.com
[6] http://smsc.cnes.fr/IASI
[7] http://www.enmap.org
[8] http://www.asi.it/en/activity/earth_observation/prisma_
[9] http://www.jspacesystems.or.jp/en_project_hisui
[10] http://hyspiri.jpl.nasa.gov

tional. Differently from the traditional multispectral (MS) sensors, hyperspectral sensors measure the solar reflected radiation in a wide wavelength spectrum (e.g., from 400 nm to 2500 nm) at narrow spectral intervals (e.g., 1 nm-10 nm). For each pixel in hyperspectral images, a near-continuous spectral signature is obtained over the whole range of wavelengths thus resulting in hundreds of bands, whereas in multispectral images it results in just few discrete spectral bands that cover some specific broad spectral wavelength ranges. An illustration of the spectral signature range (covering visible light, near infrared and middle infrared wavelength ranges) comparison between the hyperspectral EO-1 Hyperion sensor (note that the uncalibrated bands [3] are not considered) and the multispectral Landsat ETM sensor is shown in Fig.1-1. More details can be observed on the spectral signatures recorded by the Hyperion sensor. On the contrary, the multispectral Landsat ETM sensor measures only few spectral wavelength ranges, thus less details are represented in the resulting spectral signatures. This important property results in the different capability of the two types of data to describe the composition of objects of interest. However, the high number of spectral bands results in redundant information and significant noise contributions, which might affect the accuracy of the obtained results. Moreover, the high dimensionality of the hyperspectral data also leads to an increase of the computational cost. Accordingly, the general open issues in hyperspectral image analysis and processing (e.g., image classification, target detection, change detection, information retrieval, etc.) include: i) design of techniques that can effectively use the rich spectral information provided by the large number of spectral bands; ii) develop algorithms that can compress as many as possible redundant channels while preserving most of the valuable information; iii) design approaches that can be implemented in an easy yet effective way with low computational cost when exploiting the high-dimensional feature space.

TABLE 1-1 PARAMETERS OF EIGHT HYPERSPECTRAL INSTRUMENTS [2]

	AIRBORNE		**SPACEBORNE**					
Parameter	**HYDICE**	**AVIRIS**	**HYPERION**	**EnMAP**	**PRISMA**	**CHRIS**	**HyspIRI**	**IASI**
Altitude (/km)	1.6	20	705	653	614	556	626	817
Spatial resolution (/m)	0.75	20	30	30	5-30	36	60	V:1-2km H:25km
Spectral resolution (/nm)	7-14	10	10	6.5-10	10	1.3-12	4-12	0.5 cm^{-1}
Coverage (μm)	0.4-2.5	0.4-2.5	0.4-2.5	0.4-2.5	0.4-2.5	0.4-1.0	0.38-2.5 and 7.5-12	3.62-15.5 (645-2760 cm^{-1})
Number of bands	210	224	220	228	238	63	217	8461
Data cube size (samples×lines ×bands)	200×320 ×210	512×614 ×224	660×256 ×220	1000×1000 ×228	400×880 ×238	748×748 ×63	620×512 ×210	765×120 ×8461

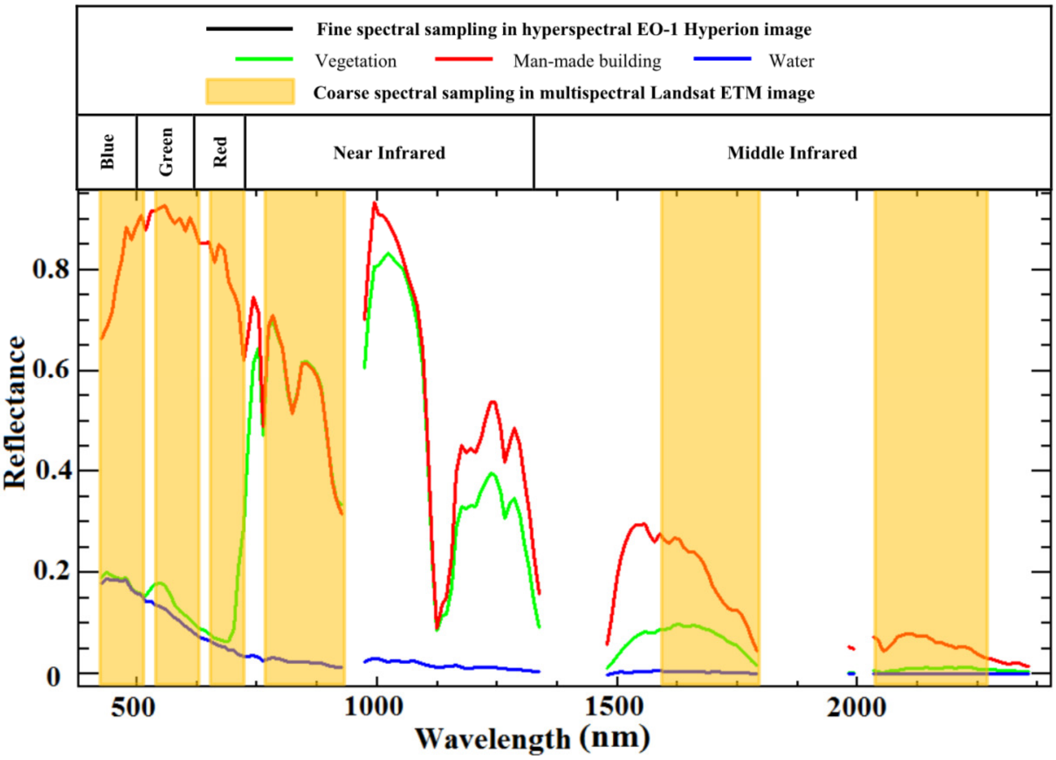


Fig.1-1 A comparison of the spectral signature ranges acquired by the hyperspectral EO-1 Hyperion sensor and the multispectral Landsat ETM sensor.

An important property of hyperspectral imaging is, that the spectral signatures of different materials have distinct spectral shapes that can be used to discriminate among each other. For example, land-cover classes: three types of vegetation, man-made building, water, which are considered in one hyperspectral image (also termed as a *hyperspectral data cube* since it has an extension along the spectral dimension) shown in Fig.1-2. A significant difference in their spectral signatures is clearly visible in the spectral domain along the wavelength. Thus one can distinguish these materials according to the unique shape of their spectral signatures (even for three kinds of vegetation that have similar spectral signatures). However, in the traditional multispectral imaging, such difference might reduce leading to the identification of only the general vegetation, man-made building and water classes. Due to the coarse spectral information represented by the discrete spectral bands in multispectral images, it is very difficult to identify the subtle classes (e.g., three vegetation classes in Fig.1-2).

The aforementioned two correlated properties of hyperspectral images (i.e., the fine spectral sampling and the discriminable spectral shapes) drive the evolution of the remote sensing image processing techniques from the multispectral into the hyperspectral images domain. In the early days' remote sensing applications, the multispectral images played a primary role due to the fact that the proposed multispectral image analysis and processing techniques were mainly based on the spatially-distributed pattern classes, thus taking advantages of spatial correlation to perform various tasks [4]. However, the hyperspectral imagery has hundreds of contiguous spectral bands allowing one to perform a more sophisticated and complex data analysis. Therefore, the target of analysis is not only those spatially distributed homogenous patterns as considered in the multispectral imagery, but also: 1) the spatially small objects which can not be simply visualized due to the limited extent of their spatial presence, but having a significant spectral difference that makes it possible to be separated from the background. This is the case of anomaly detection, man-made target detection, etc.; 2) the spectral latent variations, which only appear in some specific components of the spectral signature in hyperspectral images. As an extension from the multispectral imagery processing techniques, the hyperspectral imagery analysis should be designed considering the traditional problems in the multispectral images but also the new issues that arise due to the complex hyperspectral information. Usually a misconception is generated, which considers the hyperspectral images are just a natural extension of multispectral images due to the fact that more spectral bands are collected. This may lead to a wrong direction in addressing the hyperspectral problems just simply using multispectral processing techniques and taking advantage of the expanded spectral bands [4]. Another challenge also arises when considering the compression of the high-dimensional hyperspectral images into a low-dimensional feature space while preserving sufficient spectral information that for targets discrimination. Thus robust compression and feature extraction techniques are required to address the considered problems in hyperspectral imagery.

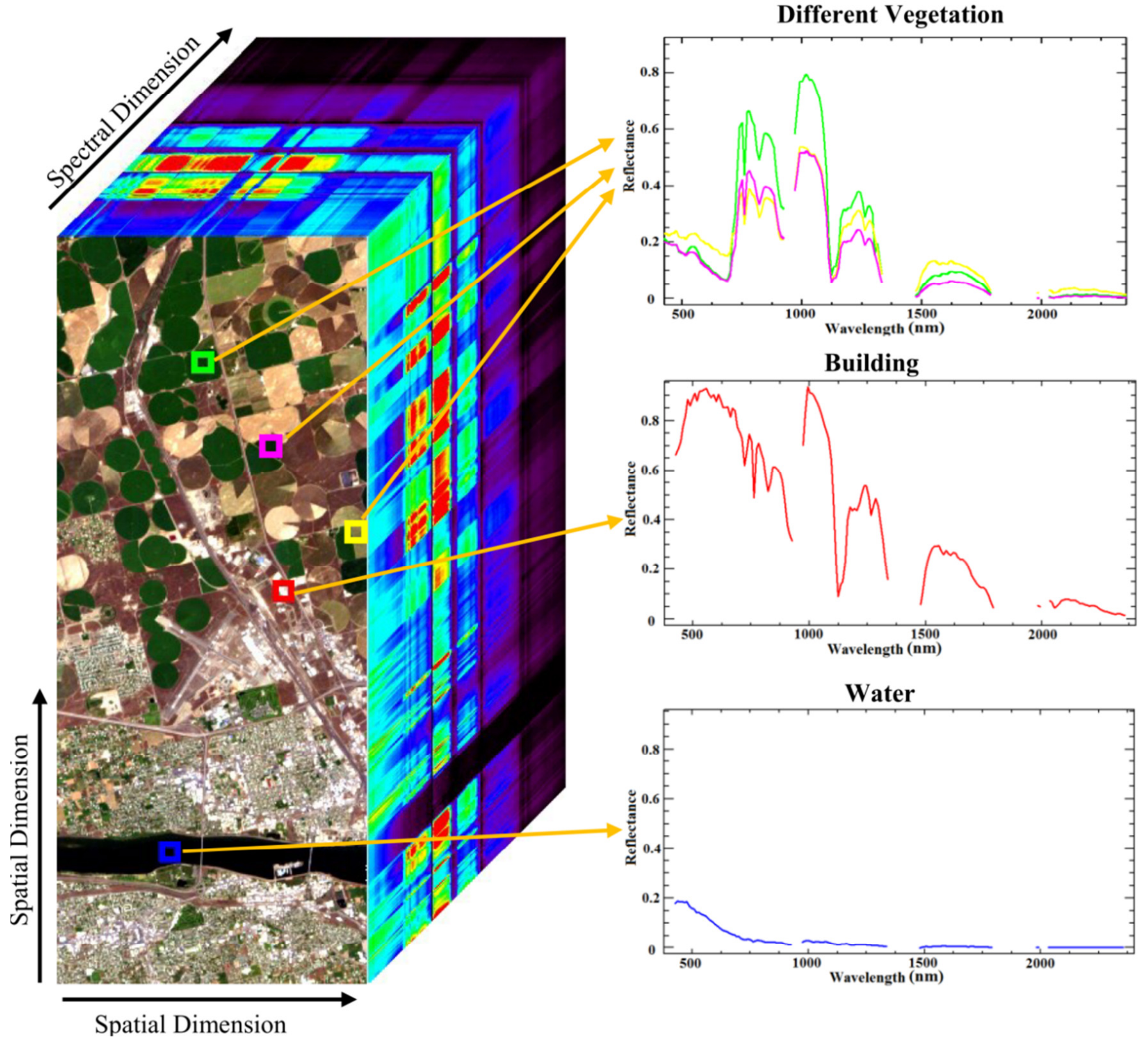


Fig.1-2 Hyperspectral data cube and examples of pixel spectral signatures associated to three land-cover materials.

An obvious consequence of the increasing data dimensionality (towards both the spatial and spectral direction) in hyperspectral images is the increase of the possible types of "targets" that can be identified. It can be explained by the phenomenon of *Heaps' law* (in information retrieval) or *Herdan's law* (in linguistics), which is an empirical law that describes the number of distinct words appeared in a document (or set of documents) as a function of the document length [5] (see a qualitative illustration in Fig.1-3). It can be approximately formulated as:

$$V_R\left(L_D\right)=w\left(L_D\right)^{\eta} \tag{1}$$

where V_R is the distinct terms in a (or some) document (s) with a length of L_D. w and η are two free parameters.

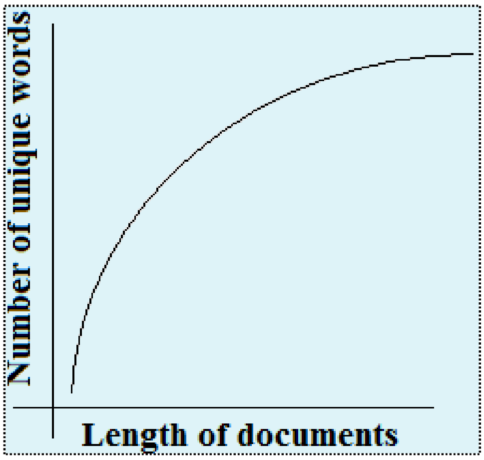


Fig.1-3 Heaps' law of vocabulary growth

The similar phenomenon occurs in the hyperspectral images, where the appeared "vocabulary" can be related to the concepts in different hyperspectral applications, for instance they can be "classes" in classification, "changes" in change detection, "small man-made objects" in target detection and "abnormal changes" in anomaly change detection field. Let $I\times J$ be the size of the considered hyperspectral image, B be the number of spectral channels. Let p be the number of "targets of interest". Thus an empirical law can be also formulated in hyperspectral images as:

$$p = H_f\left(I\times J,B\right) \tag{2}$$

where and $\mathrm{H}_f(\cdot)$ is the hidden function that describes the directly relation between the number of targets p and the dimensionality of the hyperspectral image (spatial dimension $I\times J$ and spectral dimension B). How to develop robust and advanced techniques to deal with the challenges of the increasing number of targets of interest in hyperspectral images, become a key point to guarantee an effective discovery and mining of the rich information in hyperspectral data.

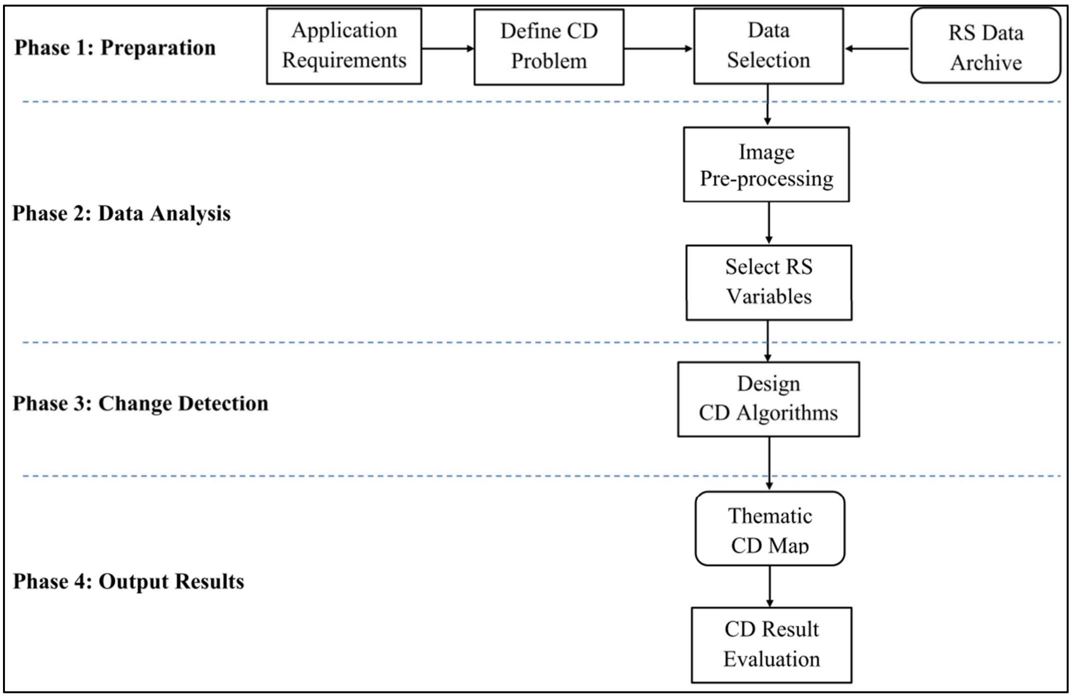
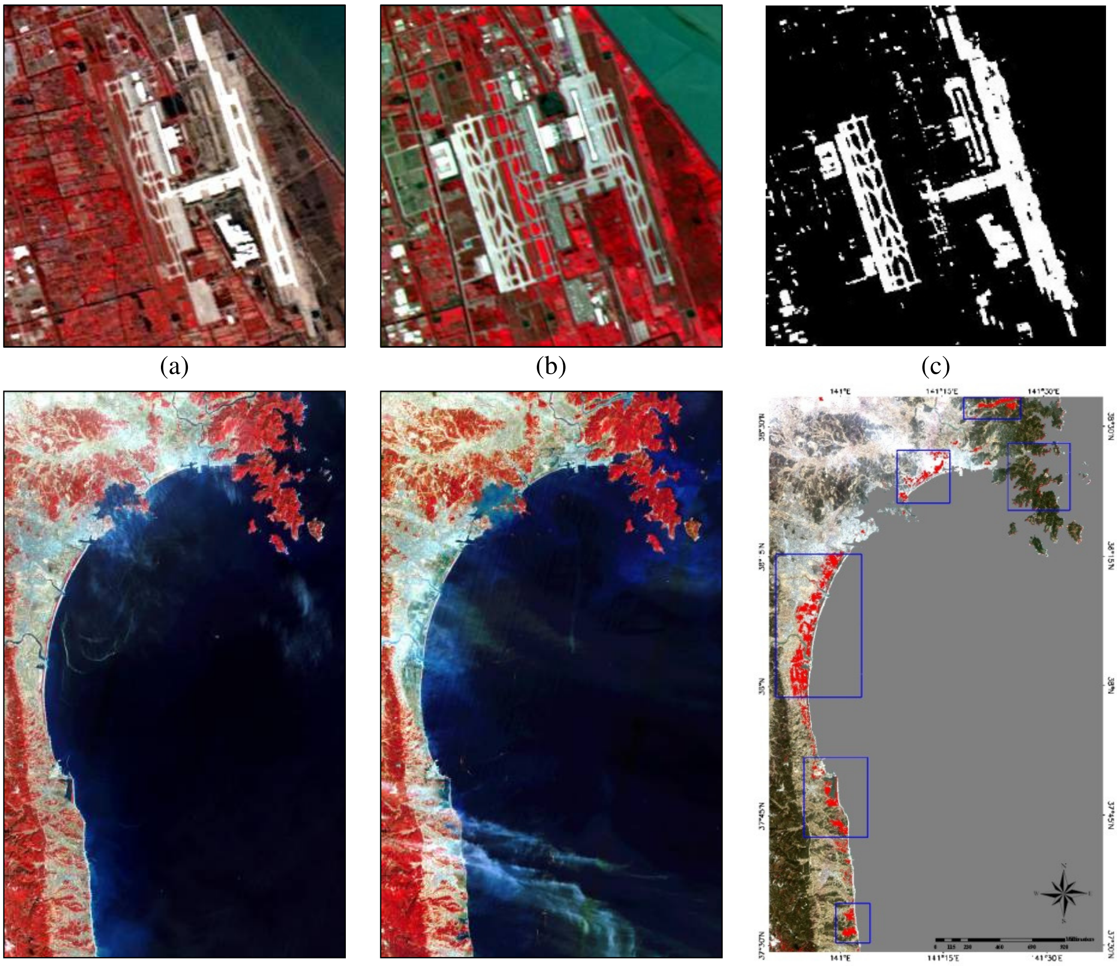


Fig.1-4 Whole change-detection processing chain.

1.2 Motivation of the Thesis

For decades, Earth Observation satellites provided a unique way to observe our living planet from space. Thanks to the revisit property of the EO satellites, a huge amount of multitemporal images are now available in archives. This allows us to monitor the land surface changes in wide geographical areas according to both long term (e.g., yearly) and short term (e.g., daily) observations. A comprehensive understanding of the global change is necessary for a sustainable development of human society. As one of the interesting subtopics in global change study, detection of anthropogenic and natural impacts on land surface is essential for environmental monitoring [6]. Change detection (CD) is one of the hottest remote sensing application topics in the past decades, which is continuously attracting attention in the remote sensing community. Technically, it is the process that identifies changes occurred between two (or more) images over a same geographical area at different observation times [6], [7], [8], [9]. Changes that reflected by the variation of image properties (e.g., pixel radiance value, texture, and shape) are mainly related to the land-cover material changes on the ground, which are the relevant changes to the real application. However, irrelevant changes may be also detected caused by some factors like variation in atmospheric conditions, sensor conditions, illumination difference and seasonal effects. The detection of the specific kinds of changes (both relevant and irrelevant changes) depends on the requirement of the real applications. It is worth noting that change detection is a comprehensive procedure (see Fig.1-4) that requires a set of processing steps [10], including: 1) understanding of the CD problem; 2) selection of suitable remote sensing data; 3) accurate image pre-processing; 4) selection of suitable remote sensing variables; 5) design of the change detection algorithm; 6) evaluation of the CD performance. In each step, effort should be devoted to drive a successful CD procedure and to result in a high CD accuracy. CD techniques have been widely and successfully used in several remote sensing applications (e.g., ecosystem evolution, urban area study, disaster monitoring) in the past decades, especially considering the multitemporal multispectral images [6], [8], [11]. Two examples of change detection applications are shown in Fig.1-5 (a) and (b): monitoring of the urban infrastructure (i.e., international airport) construction [12], and monitoring of the damaged areas in natural disaster (i.e., tsunami) [13].



(d) (e) (f)

Fig.1-5 Examples of two change-detection applications. First column, a monitor of the construction of Shanghai PuDong International Airport by using bi-temporal CBERS satellite images acquired on (a) March, 7, 2005 and (b) May 7, 2009. (c) the obtained CD map [12]; Second column, a fast monitor "3.11" Japan earthquake-triggered tsunami disaster in 2011 from HJ-1A/B satellite images acquired on (d) February 24 (before tsunami), and (e) March 14 (3 days after tsunami), (c) the obtained CD map [13].

In general, CD techniques aim at automatically detecting changes occurred between multitemporal images. Thus by using these techniques one can gradually reduce the need for conventional field investigations in real CD applications. This is the main motivation of the CD techniques. Due to the coarse spectral sampling in a few discrete spectral bands and the insufficient spectral representation in multispectral images, the early stage of the development of the change detection techniques in multispectral images (CD-MS) is mainly focused on the abrupt changes. They are land-cover class transitions that significantly affect the spectral signature (e.g., vegetation to land covers like water, built-up areas, soil). For change detection in hyperspectral images (CD-HS), by taking advantage of the detailed spectral sampling, the aim is to detect the spatially homogenous and spectrally significant changes (at global scale) associated to the land-cover class transitions as in CD-MS, but also the spectrally insignificant subtle changes (at local scale), which are usually not detectable when employing multispectral images. These changes usually locate in a (some) portions of the whole spectral signature. In order to effectively perform CD-HS and obtain highly accurate results, it is important to design advanced CD techniques that can take advantages of the properties of the multitemporal hyperspectral images acquired by the new generation of hyperspectral remote sensing satellite systems, and properly analyse and identify variations in the spectral-temporal domain.

In addition to the main motivation arising by the hyperspectral data properties, some other important points also drive the motivation of this thesis on change detection techniques:

From the availability of the ground truth data point of view, CD techniques can be split into two main groups: supervised and unsupervised. The supervised CD techniques are based on supervised classification schemes and assume that ground truth or prior knowledge is available for the training of a classifier. Although supervised CD methods generally outperform the unsupervised ones in detecting land cover transitions with high accuracy, the process of collecting reference data for multitemporal images is time consuming and costly, and often unfeasible. Therefore, unsupervised CD approaches that do not rely on any reference sample are more attractive from the practical point of view. In this thesis, the design of the CD approaches is mainly in a partially-unsupervised or unsupervised fashion.

From the application point of view, two families of the CD techniques can be identified: 1) binary change detection techniques or 2) multiple-change detection techniques. Binary CD aims at detecting the presence/absence of change without giving any information about the possible separation of multiple changes. Thus, all kinds of changes present are simply considered as one single general change class. For multiple change detection, the aim is not only to detect the changes, but also to identify different kinds of changes among each other. Since few literature works are devoted to solve the multiple-change detection problem in multitemporal hyperspectral images, this thesis focuses on this interesting but challenging topic while investigating in details the model of the problem and the analysis of the changes.

1.3 Objectives of the Thesis

The high spectral resolution and narrow spectral intervals directly lead to an increase in the data dimensionality, as well as to the presence of redundant information in hyperspectral images. This makes the change analysis more complex and challenging. In reality, the existing CD approaches are mainly designed for multitemporal multispectral images, which efficiency is poor when directly applied to hyperspectral images. In this thesis, the main objective is to define advanced CD techniques to solve the multiple-change detection problem in multitemporal hyperspectral images, in order to meet the requirements of practical CD-HS applications especially when the ground truth data are not available. In particular, the following issues are investigated in detail in the thesis:

1) The proper definition of the multiple-change detection problem in multitemporal hyperspectral images, and the analysis of the change structure and the relation among changes;

2) The investigation of reliable change index (or domain) for representing and analysing multiple changes in the high-dimensional or compressed low-dimensional feature spaces;

3) The design of techniques for effectively discovering, visualizing, and detecting multiple changes. Development of a user-friendly CD tool that allows users to have an easy yet efficient implementation;

4) The design of automatic (or semi-automatic) CD techniques for the detection and separation of multiple changes according to the clustering nature of spectral signatures;

5) The investigation of spectral variations at sub-pixel level, thus to exploit in detail the possible kinds of changes inside a pixel to make better decision for change identification;

6) The design of advanced unsupervised and automatic CD techniques which are independent from the availability of the ground truth data.

1.4 Novel Contributions of the Thesis

Based on the main motivation and objective of the thesis, attention is focused on the development of the advanced techniques for automatic change detection in multitemporal hyperspectral images. Research activities are mainly carried out to develop robust techniques for addressing the considered challenging CD-HS problem. The main contributions and novelties of the thesis are briefly reported as follows.

i) A theoretical and empirical analysis of the considered CD-HS problem in the spectral difference domain and multitemporal spectral stacked domain

By taking into account the intrinsic complexity of the hyperspectral data, a proper definition of the concept of changes in hyperspectral images is given in the spectral difference domain, which is computed by subtracting the multitemporal images pixel by pixel. Two kinds of changes are defined with respect to their levels of spectral change significance: the major changes and the subtle images. A hierarchical nature of the spectral changes is observed by analysing in detail the spectral variations from coarse to fine processing levels, leading to a better modelling of the hidden and complex change structures. This analysis is exploited in (ii) and (iii) (see below).

Another interesting and important analysis for modelling the multiple change detection problem is proposed in the multitemporal spectral stacked domain, which provides a new perspective to detect changes by jointly exploiting the spectral-temporal variations at subpixel level. The main advantages of working in the spectral stacked domain are: 1) it preserves the intrinsic properties of the spectral signatures that represent the real land-cover materials, which are extended along the temporal direction; 2) only the occurred land-cover transitions are identified as endmembers in the mixture model, which usually are generated and considered when implementing unmixing independently on single-time images (even do not exist). These are two fundamentals that support the proposed multitemporal spectral unmixing based CD approach described in (iv) (see below).

ii) A novel semi-automatic sequential spectral change vector analysis (S^2CVA) for discovering, representing and detecting multiple changes

S^2CVA is developed based on the state-of-the-art C^2VA approach. The proposed approach aims at discovering, representing and detecting multiple changes according to a sequential process that takes into account different levels of spectral change significance. The main novelties of the proposed S^2CVA are: 1) it iteratively analyzes the heterogeneous change information by following a top-down structure and a sequential analysis. The change information in the original high dimensional feature space are adaptively and iteratively compressed and projected into new 2-D feature spaces, each of which is associated to a specific portion of the whole spectral change vectors. Thus changes can be represented, discovered and detected at different levels of the hierarchy; 2) at each level it adaptively exploits and represents multiple-change information is a 2-D change representation domain by automatically selecting a reference vector that maximizes the measurement of data variance.

iii) A novel partially-unsupervised hierarchical clustering method for multitemporal hyperspectral images change detection

The main contributions in this work are as follows: 1) proposal of a technique for addressing the challenging multiple-change detection problem in hyperspectral images, by considering the difference of spectral change behaviours in the spectral difference domain at different spectral scales; and 2) proposal of an approach that models the detection of multiple changes in a hierarchical way, to identify the change information and separate different kinds of changes (major change, subtle change, and finally, change endmembers) according to the spectral homogeneity. By this way, the complex CD-HS problem is progressively decomposed into several specific sub-problems, focusing on each single portion of the multiple-change information. This makes it possible to discover the difference among similar changes by decreasing the difficulty of detection. Moreover, the proposed approach is designed in a partially-unsupervised way, where a manual initialization can be easily implemented to trigger an automatic model selection and clustering at each level.

iv) A novel unsupervised Multitemporal Spectral Unmixing (MSU) for detecting multiple changes in hyperspectral images

The proposed MSU approach is designed in a fully automatic and unsupervised way, thus is independent from the availability of prior knowledge and the manual assistance of the user in the real applications. The main novelties and contributions of the proposed method are as follows: 1) it provides a new perspective to detect changes by jointly exploring the spectral-temporal variations in the multitemporal spectral stacked domain; 2) it proposes a multitemporal spectral unmixing framework to solve the multiple change detection problem, where the identification of the number of the change classes is done by identifying the distinct endmembers and the unique change classes, and the discrimination of changes is addressed by unmixing and abundances analysis; 3) it allows one to understand in details the spectral composition of a pixel, thus implementing CD at subpixel level, whereas to our knowledge that most of the state-of-the-art techniques are designed at pixel level only. By taking advantage of the endmembers extraction and spectral unmixing while considering endmember variability (i.e., local endmembers strategy), the proposed MSU method well models the change and no-change spectral compositions inside a pixel. A more reliable decision is made according to the analysis of the endmember abundances associated with a given class with respect to a crisp decision based on the pure-pixel theory. Accordingly, more subpixel level spectral variations are expected to be identified, which are usually not detectable in the pixel-level-based CD techniques.

1.5 Structure of the Thesis

This thesis is organized in seven chapters. The present chapter gives a brief introduction on the remote sensing and the new generation of the hyperspectral sensors. It presents in details the motivation of the thesis on change-detection techniques for multitemporal hyperspectral remote sensing images. Then the main objectives of the thesis are introduced. The novel contributions of the thesis are given with a brief summary on each of them. Finally, it describes the structure of the whole thesis.

Chapter 2 presents an intensive review of the state-of-the-art change detection techniques in multitemporal multispectral and hyperspectral images, respectively. Problems and challenges that arise when changing the perspective of CD from multispectral to hyperspectral images are analyzed and discussed in details.

Chapter 3 introduces several important and novel concepts for multiple-change detection in hyperspectral images. The spectral difference domain and the multitemporal spectral stacked domain are analyzed in details. This analysis resulted in the design of the advanced CD techniques that proposed and presented in the next chapters.

Chapter 4 presents a novel semi-automatic sequential spectral change vector analysis for discovering, representing and detecting multiple kinds of changes in multitemporal hyperspectral images. The proposed approach provides also an easy yet effective tool for user-interaction within the CD procedure.

Chapter 5 introduces a novel partially-supervised hierarchical clustering method for multitemporal hyperspectral images change detection. The proposed approach is developed following a hierarchical top-down structure, and a manual initialization and adaptive clustering is included at each iteration to exploit the number of the hidden changes and to separate them.

Chapter 6 describes a novel automatic multitemporal spectral unmixing approach to address the multiple-change detection problem in hyperspectral images. The proposed approach is designed based on a spectral multitemporal unmixing technique at sub-pixel level, thus is able to investigate in details the spectral-temporal variations within a pixel.

Chapter 7 draws the conclusions of this thesis. The remaining open issues and further developments of the research activities are also discussed.

Chapter 2

State-of-the-Art: Change Detection Techniques

This chapter gives a comprehensive overview of change-detection techniques presented in the literatures for both multispectral and hyperspectral multitemporal images. Problems and challenges that may arise due to the change of perspective from multispectral to hyperspectral images are analyzed and discussed in detail.

2.1 Change Detection Techniques for Multitemporal Multispectral Images

For decades, many CD techniques have been proposed for successfully addressing the CD problem in multispectral optical remote sensing images. Some good reviews for these CD techniques can be found in [6], [7], [8], [9], [14], [15], [16], [10]. In this chapter, we focus on the overview of the techniques for the main change-detection step in the whole CD processing chain (i.e., Fig.1-4, Phase 3). In the rest of this section, an overview on the relevant literature is present.

As mentioned in the Introduction, from the methodological point of view, CD techniques can be clustered into supervised and unsupervised depending if they need ground truth data or not. Supervised CD methods are mainly designed based on the supervised classification schemes and require the available prior knowledge for training a classifier. In this context, the most popular CD approach is the Post-Classification Comparison (PCC) [17], [18], which classifies independently two (or more) images at different times and then compares the pixel class label to detect changes. The main advantages of PCC is that the land-cover transitions are obtained (i.e., "from-to" information). However, the accuracy of the change-detection performance highly depends on the accuracy of the classifier. Moreover, the classification errors on each single date image impact on the final change-detection accuracy. Another type of supervised CD approaches is based on the Direct Multi-date Classification (DMC) [19], [20], [21], [22]. It identifies changes by simultaneously classifying the stacked multi-date images, thus a change is represented by an output class in the final classification map. The considered CD task actually is replaced by a classification task. However, difficulties come from the generation of comprehensive multitemporal training samples that represent the detailed land-cover transitions between the multitemporal images, which might highly affect the classification accuracy. In reality, it is very challenging (in some cases impossible) to have such fine multitemporal training samples available. In addition, another group of supervised CD approach is based on analysis and classification of the change index images instead of working on the original images (i.e., spectral channels). For example, in [23], the binary Support Vector Machine (SVM) was applied to the spectral difference images stacked with different extracted features to detect the land-cover changes in the mining area; in [24], SVM was also used for solving a binary CD problem for monitoring urban growth by classifying an improved change index image fused from various spectral difference images (e.g., differencing, ratioing, distance metric, similarity measure). Note that such supervised CD methods are mainly developed for multispectral images but also are applicable to hyperspectral images as well. However, when dealing with hyperspectral data, more attention should be devoted to define effective classification systems that: i) are suitable to the analysis of high-dimensional data and overcome the Huges phenomenon (i.e., with a fixed number of training samples, the predictive power of a classifier reduces as the dimensionality increases) [25], and ii) can effectively exploit informative features thus enhancing change detectability. Although supervised CD methods generally outperform the unsupervised ones in detecting land-cover transitions, the process of collecting reference data for multitemporal images is always time costly and often unfeasible. To overcome this drawback, some works were designed when a small portion of the reference samples are available [26], [27], which is called as partially-unsupervised or semi-supervised learning. The main idea is to start from an unsupervised procedure to create the initial training samples for a supervised or semi-supervised classification strategy, or to start with some available initial samples to learn and add more informative samples in a supervised classification thus to enhance the CD performance.

Despite the effectiveness and usefulness of all these supervised and semi-supervised CD approaches, unsupervised methods that do not rely on any ground truth data or prior knowledge are more attractive from the real application point of view. Unsupervised CD methods have been designed for both binary change detection (i.e., considers only the presence/absence of change, ignoring the possible different class transitions) and multiple-change detection (i.e., detects the changes, but also identifies their difference among each other). Binary change detection only distinguishes between change and no-change classes and it can not provide detailed information about the class transitions. Thus it is usually just an initial understanding of the spatial distribution of the changes on the images in a satellite observation period. From the methodological point of view, they can be categorized into thresholding-based and clustering-based techniques. The thresholding-based techniques are designed to find a proper threshold value that separates the change and no-change two classes on the bi-modal histogram of the magnitude image. To this end, some classical image segmentation approaches can be used, for example, OSTU algorithm [28], Kittler-Illingworth (K-I) algorithm [29], and maximum entropy thresholding [30], etc. In [31], the problem of binary CD was solved automatically by modeling the statistical distribution of classes as Gaussian and incorporating spatial-context information, thus significantly improved the previous works that are mainly based on manual thresholding [32]. Statistical distributions were theoretically discussed in [33] thus to better model the mixture of change and no-change two classes in the magnitude domain. Clustering algorithms have been also investigated and used for solving the same binary CD problem as well, for instance, the Principal Component Analysis (PCA) and *k*-means clustering were combined in [34]. In [35], Fuzzy c-means (FCM) and Gustafson-Kessel Clustering (GKC) algorithms were applied with two other optimization techniques, Genetic Algorithm (GA) and Simulated Annealing (SA) to further enhance the CD performance. In [36], a nonlinear support vector clustering was designed for separating the change and no-change binary information. A fuzzy clustering algorithm with a modified Markov Random Fields (MRF) energy function was proposed in [37], and an unsupervised CD approaches was designed in [38] by mapping the difference image in the feature space and applying kernel *k*-means clustering.

However, in multispectral images the multiple changes can be detected. Usually two main issues have to be considered simultaneously: 1) to correctly identify the number of multiple changes; 2) to effectively discriminate different changes among each other. In literature, many attempts based on image transformation, multivariate analysis, etc., have been done to address the multiple-change detection problem. The most popular approaches are for example the one proposed in [39] by applying PCA on the difference image and analyzing the changes that characterized in the first few PCs. In [40], the Gramm-Schmidt (GS) transformation was used. In [41], the Tasselled cap transformation (KT) was applied to detect the vegetation change from Landsat TM images. In [42], an unsupervised Multivariate Alteration Detection (MAD) technique based on the Canonical Correlation Analysis (CCA) was used to detect the seasonal vegetation changes. An improved version named Iterative Reweighted MAD (IR-MAD) was proposed in [43] to provide more reliable output components thus emphasizing changes. However, the main disadvantages of the above mentioned transformation-based CD approaches is that they require a strong interaction with the end-users to select the most informative components thus to emphasize on the specific changes, which is usually time consuming and application-oriented. On the other hand, the transformation-based methods do not provide a clear number of changes. The number of detected changes highly depends on the selected number of components and the change information represented in those components. Some changes might be still mixed and unidentified in a given component. Therefore, the transformation-based approaches are good at extracting features for enhancing the detection of specific kinds of changes in CD, but in general not suitable for detecting all the possible change classes.

Another popular and classical method for multiple change detection is Change Vector Analysis (CVA). It was proposed in 1980 by Malila [44]. CVA models a change vector by a direction and a magnitude. Different kinds of changes can be identified by analyzing these two variables. Many works were developed based on CVA to extend its use in different applications and proposed its improved versions [45], [46]. In [33], Bovolo et al. proposed a framework for a formal definition to the CVA in a polar coordinates. In such polar representation domain, the spectral change vectors are represented and distributed according to their intrinsic properties and reveal the nature of changed and unchanged pixels. This work also provided a solid background that for developing more advanced and accurate automatic change detection algorithms. However, the CVA method needs user to select 2 out of *B* spectral bands at each implementation thus to discover specific changes of interest. To overcome this limitation, a simple but effective method named Compressed Change Vector Analysis (C^2VA) was recently proposed in [47]. In C^2VA the considered multiple-change detection problem is represented in a magnitude-direction 2-Dimensional polar domain generated by a lossy compression (potentially ambiguous) procedure on the original *B*-Dimensional feature space. Thus the change detection can be easily implemented without relying on any band selection, which might result in loss of change information. C^2VA has proved to be effective in different CD applications with multispectral images [47], [48]. Differently from the aforementioned transformation-based CD approaches, the uncompressed CVA and compressed C^2VA provide an opportunity to understand the changes that represented in the selected or the whole spectral difference feature space. Thus the number of changes and the discrimination of different changes can be exploited by analyzing the corresponding representations.

Several other works exist devoted to improve the CD result by considering the spatial information in the CD procedure thus reduce the "salt and pepper" noise caused by the pixel-based processing. In this context, two main groups of methods are reviewed. The first large group of methods jointly uses spectral-spatial information to improve the CD performance. Spatial information extraction and representation techniques like spatial neighborhood features [49], Gabor filters [50], Markov Random Fields (MRFs) [31], Morphological Attribute Profiles (MAPs) [51], fusion of textures, edges and others spatial features [52], etc. were developed in the literature. The proper and effective use of the extracted spatial features is the key point to ensure a good CD result. Another group is the object-oriented CD approaches, which usually designed based on the image segmentation and change object analysis. For instance, in [53] CD approaches were proposed based on object/neighborhood correlation image analysis and image segmentation techniques. A statistical object-based method was designed based on segmentation, image differencing and stochastic analysis of the multispectral signals [54]. A parcel-based context-sensitive technique was investigated to improve the pixel-based CD performance in [55], where a multilevel CVA was applied. In [56], an object-oriented CD algorithm was developed by analyzing the abnormal statistical properties of the segmented objects in ArcGIS. Either the spatial feature based or the object-oriented methods are proven to be effective in obtaining more spatially homogenous CD result, reducing a lot of the detection errors, especially the commission errors.

Other works also focus on the improving of the CD results by using different new techniques that take advantages of the developments in the machine learning and pattern recognition society. For example, in [57] an approach was proposed for binary CD by using the optimized computation algorithm (i.e., genetic algorithm). In [12], multiple difference index images were combined to construct better change variables for binary CD using different ensemble learning schemes. In [58], different levels of data fusion approaches were analyzed, and a sequential fusion based CD procedure were proposed to utilize information from multi-sensor images over the same geographical area. In [52], fusion strategies for multiple features (e.g., texture, shape, edge) extracted from the original images were developed to improve the CD result which usually only rely on the spectral information; A Slow Feature Analysis (SFA) algorithm was proposed for multispectral images change detection in [59], which can extract the most temporally invariant components from the multitemporal data to construct a new feature space for separating the change and no-change. A clustering approach based on a semi-nonnegative matrix factorization (semi-NMF) was proposed in [60], which proven to be simple in computation yet effective in identifying meaningful changes. A sparse hierarchical clustering approach was proposed in [61]. Discriminative change features were generated by stacking bitemporal multi-scale center-symmetric local binary pattern features. Then a tree-structured dictionary was built and the sparse reconstruction error was used to model the changes. Despite these proposed CD techniques have shown their advantages in different specific cases in literature works, the effectiveness and applicability of these approaches strongly depends on the considered data sets and the complexity of the real CD tasks. In general, it is difficult to find a unified approach that suitable for all CD cases.

2.2 Change Detection Techniques for Multitemporal Hyperspectral Images

Change detection techniques in multispectral images have been exhaustively investigated in the past few decades. However, there are relatively few works on CD in hyperspectral images, rarely are for solving the multiple-change detection problem. These works can be divided into three main categories including: 1) transformation based methods; 2) spectrum analysis based methods, and 3) other techniques.

In the first category, cross covariance (chronochrome) and Covariance-Equalisation (CE) are two multivariate statistical techniques, which detect differences between linear combination of the spectral bands (i.e., subtracting feature vectors in the transformed feature space) from the two acquisitions. They have been applied in CD in [62], [63]; In [64], three iterative clustering methods: class-conditional CE (QCE), bitemporal QCE and Wavelength Dependent Segmentation (WDS), were applied to detect man-made changes in VNIR and TIR hyperspectral images. It proved that the use of a spatially adaptive detector greatly enhance the CD performance for both target detection and false alarm reduction. Another popular transform-based method represents the images in a feature space, where the change information is concentrated in few components. This reduces the data dimensionality and noise, and focuses on the components that are related to the specific changes of interest. In this context, Multivariate Alteration Detection (MAD) method was introduced in [42] to solve a vegetation CD problem by using multitemporal hyperspectral images in an unsupervised way. Then it was extended to an iterative reweighted version (IR-MAD) [43] to provide more reliable output components and thus better emphasizes and detects changes. In [65], two kernel versions of Maximum Autocorrelation Factor (MAF) analysis and Minimum Noise Fraction (MNF) analysis were introduced for CD. The experimental results showed that the kernel MAF/MNF performed better than its linear version and the kernel PCA. A Temporal-Principal Component Analysis (T-PCA) was proposed in [66], which exploits the variances in PCs after transforming the combined multitemporal images. Thus the no-change and change information is associated with the first and second group of PCs, respectively. Another variation of PCA named sparse PCA was also investigated recently in [67] for addressing the CD problem, and the result showed that three wavelength regions were important for CD, which might be also useful for the potential feature selection. In [68], Independent Component Analysis (ICA) was applied with the Uniform Feature Design (UFD) strategy in a hierarchically framework to investigate the specific vegetation changes. In [69], a spatio-temporal ICA (stICA) was designed for extracting the spatio-temporal patterns from different hyperspectral sensors or from different acquisition conditions and dates. It is worth noting that the transformed-based methods require a strong interaction with the end-users to select the most informative components thus to emphasize on specific changes. This step is usually time-consuming, especially when the number of changes is large in the hyperspectral case. Therefore, it is not suitable to detect all possible change classes. Moreover, transformation based approaches are not able to provide a reliable number of changes, which limited the use of these methods in the real CD-HS applications.

The spectrum analysis based CD methods take advantage of the detailed spectral signature in hyperspectral images. On the one hand, the distance and similarity measurements are used to detect the difference between the considered pixel spectral signatures of images acquired at two times (e.g., Spectral Angle Measure (SAM), Spectral Information Divergence (SID) and Spectral Correlation Measure (SCM) [70], [71], [72]. On the other hand, CD techniques were designed based on the spectral properties of the hyperspectral images. In [73], a linear mixture model was proposed for analyzing the endmembers and abundances estimated from each single-time image to address a binary CD problem. In [74], a simple but effective relative radiometric normalization method was analyzed and two automatic approaches for CD-HS after normalization were introduced. A novel multi-spectro-temporal analysis approach was proposed in [75] based on 3-D spectral modeling and multi-linear algebra, thus to model the temporal variation of the reflectance response as a function of time period and wavelength. In [76], a subspace-based CD method was designed when knowing the undesired land-cover type spectral signature as a prior knowledge. The subspace distance was computed to determine whether the target was anomalous with respect to the background subspace. Those anomalous pixels were then considered as changes.

Other works have been also developed to explore the CD-HS problem from different perspectives. A CD approach based on tensor-factorization and PCA was proposed in [77], which analyzed the concatenated multitemporal hyperspectral image as a 3-Dimensional tensor cube. A model-based methods by formulating the CD as a statistical hypothesis test was presented in [78], and its application to Airborne VNIR/SWIR hyperspectral images was discussed in [79]. An unsupervised approach was developed in [80] to achieve slight change extraction and detection in hyperspectral images, which also exploits the statistical inner connection among the multitemporal image sequence. A semi-supervised CD method was proposed in [81] by designing a new distance metric learning framework for CD in noisy conditions. Moreover, there are also some other works focusing on the external factors that affect the CD performance, which include limiting image parallax errors [82], vegetation and illumination variation [83], and diurnal and seasonal variations [84]. These factors may introduce errors into the CD process and thus decrease the detection accuracy, which should be limited as much as possible in real CD-HS applications.

2.3 Problems and Challenges

Due to the intrinsic properties of the hyperspectral data, the CD problem becomes more complex and challenging thus efficiency of the exist methods (especially the one designed for multispectral images) is reduced. The main challenges are analyzed and summarized as follows:

a) High dimensionality of hyperspectral images. It involves challenges in data handling, including storage volume and computing bottle necks, which are actually common problems for all hyperspectral data processing tasks (i.e., classification, change detection, target recognition). In CD-HS, the main difficulty is to effectively extract the change information from the high dimensional feature space, which makes the changes more implicit and less separable. Techniques that developed for multispectral images, like the standard CVA or C^2VA [33], [44], [47], may fail to give a proper change representation and effective change identification, thus degrade their performance and decrease the CD accuracy. This is mainly due to the ambiguity that generated by the compressed representation, and also to the potentially critical situation when too many changes are present. Moreover, difficulty also comes from the high computational cost for change representation and identification in the high dimensional feature space, which may limit the use of many CD techniques.
b) Rich but complex spectral variation information. The fine spectral sampling in hyperspectral images leads to many possible subtle changes usually undetectable in multispectral case. However, subtle changes are highly difficult to be identified, as they follow a complex structure in the high-dimensional feature space. Variations can be characterized in a specific part of the whole spectrum thus are not easy to be identified from a global point of view. The existing CD-HS methods try to extract all changes directly from the original data space or from a transformed feature space relying only on a single level analysis, which increases the difficulties of identifying multiple-change classes and thus affects the detection accuracy. Moreover, most of the existing unsupervised CD methods that directly compare and analyze the difference of pixel radiance values, ignoring the rich near-continuous spectrum information that is the peculiar property of hyperspectral images, might be ineffective to detect the subtle spectral variations.
c) Redundant information. The spectral information of the adjacent bands in hyperspectral images results in a non-negligible redundancy. However, when the spectral resolution increases, a reduction of the signal-to-noise ratio (SNR) of the spectral signal is obtained [85]. Thus information represented in a single hyperspectral band becomes more sparse and implicit, which may degrade the discriminability of a detector.
d) Most of CD-HS approaches present in the literature focus on either binary CD (e.g., [73], [74], [77], [80], [81]) or the detection of specific changes (e.g., [42], [64], [65], [68], [70], [76], [78]). There are few methods that address the challenging problem of detecting multiple changes simultaneously (which can be very important especially when unexpected changes occur on the ground). Moreover, some methods still rely solely on the change magnitude information [42], [43], [68], neglecting the very interesting change direction information (and thus the whole spectrum signature) for distinguishing different kinds of changes. Although the transformation-based methods (e.g., MAD, IR-MAD, TPCA, kernel MAF/MNF, ICA) allow one to detect multiple changes, the application of transformation directly to the complex hyperspectral data results in: 1) a high computational cost; 2) a difficult manual interpretation of all the transformed components thus select the ones related to the chang-

es of interest; and 3) a qualitative and ambiguous description of change classes, especially for the subtle and latent changes. Therefore, the identification of the number of multiple changes and the discrimination among them from the transformed components are still open issues.

Chapter 3

Proposed Concepts for Change Detection in Hyperspectral Images

In this chapter, some proposed new concepts for multiple-change detection in multitemporal hyperspectral images are presented. First of all, an analysis of the spectral difference domain is provided. By taking into account the intrinsic complexity of the hyperspectral data, some concepts associated to the multiple-changes are given from the perspective of pixel spectral signature in order to formalize the considered CD problem. Then, a multitemporal spectral stacked domain is introduced and analyzed. A multitemporal spectral mixture model is defined to investigate in detail the spectral composition of the change and no-change classes within a pixel. Thus the considered multiple-change detection problem can be solved by the proposed multitemporal spectral unmixing technique.

3.1 Analysis of the Spectral Difference Domain

The aim of this section is to investigate the spectral difference domain and its role in addressing the multiple-change detection problem in hyperspectral images. It is necessary to identify the class transitions having discriminable spectral behaviors either globally or locally in the spectrum of multitemporal hyperspectral images. Based on the pure pixel assumption that assumes a pixel contains only a kind of material substance inside, a *Spectral Change Vector* (SCV) in the spectral difference domain is assumed to associate to a given land-cover class transition. Then the discriminable SCVs are defined here as *change endmembers*. The clusters of the change endmembers are analogous to the change classes that we are detecting in the multitemporal images.

Note that the very high spectral resolution in hyperspectral images makes it possible to detect many differences in the spectral signatures of pixels acquired in a scene of interest. Such differences may occur at different spectral resolution levels. In order to conduct an effective CD-HS, it is important to understand and model the concept of "change" in the multitemporal hyperspectral images and its relationship with the concept of change endmembers.

In the change detection research, the most popular change index image is the spectral difference image $\boldsymbol{X}_D$ (or named SCVs in a B-Dimensional spectral difference domain $\boldsymbol{X}_D$). Numerous works have been done by analyzing the difference image $\boldsymbol{X}_D$ to detect changes [6], [12], [31], [33], [47], [86]. Let $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ be two co-registered hyperspectral images having a size of $I \times J$ acquired over the same geographical area at

---

Parts of contents in this chapter are taken from:
Book Chapter: submitted to Springer EARSeL Book 'Multitemporal Remote Sensing: Method and Applications'. Chapter title: "Change Detection in Multitemporal Hyperspectral Images". Authors: L.Bruzzone, S.Liu, F.Bovolo, P.Du.
Journal Paper: submitted to IEEE Transactions on Geoscience and Remote Sensing. Title: "Unsupervised Multitemporal Spectral Unmixing for Detecting Land-Cover Transitions in Hyperspectral Images". Authors: S.Liu, L.Bruzzone, F.Bovolo, P.Du.

times t_1 and t_2, respectively. Let $x_1(i, j)$ and $x_2(i, j)$ be the pixels with spatial position (i, j) ($1 \le i \le I$, $1 \le j \le J$) in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively. To analyze the behaviors of spectral differences between the two images, let us compute the hyperspectral difference image $\boldsymbol{X}_D$ (and thus the SCVs associated with each pixel) by subtracting bitemporal images from each other pixel by pixel [33].

$$\boldsymbol{X}_D = \boldsymbol{X}_2 - \boldsymbol{X}_1 \tag{3}$$

In $\boldsymbol{X}_D$, each pixel is characterized by a SCV that shows as many elements as the spectral channels in the original hyperspectral images. Each element assumes values that depend on whether a change occurred or not for a specific wavelength, and on the kinds of changes. Therefore, SCV signatures that are related to the land-cover class transitions in $\boldsymbol{X}_D$ are used to formalize the considered CD problem.

Let $\Omega = \{\Omega_n, \Omega_c\}$ be the set of all classes in $\boldsymbol{X}_D$, where Ω_n is the set of no-change class and $\Omega_c = \left\{\omega_{c_1}, \omega_{c_2}, ..., \omega_{c_K}\right\}$ is the set of the K possible change classes. Note that since we are not interest in distinguishing different kinds of no-change in Ω_n, thus Ω_n is considered as one general no-change class ω_n, $\Omega_n \approx \omega_n$. So the considered multiple-change detection problem can be defined as to detect all changed pixels (Ω_c) in $\Omega = \{\omega_n, \Omega_c\}$, and to separate them into multiple change classes $\left\{\omega_{c_1}, \omega_{c_2}, ..., \omega_{c_K}\right\}$.

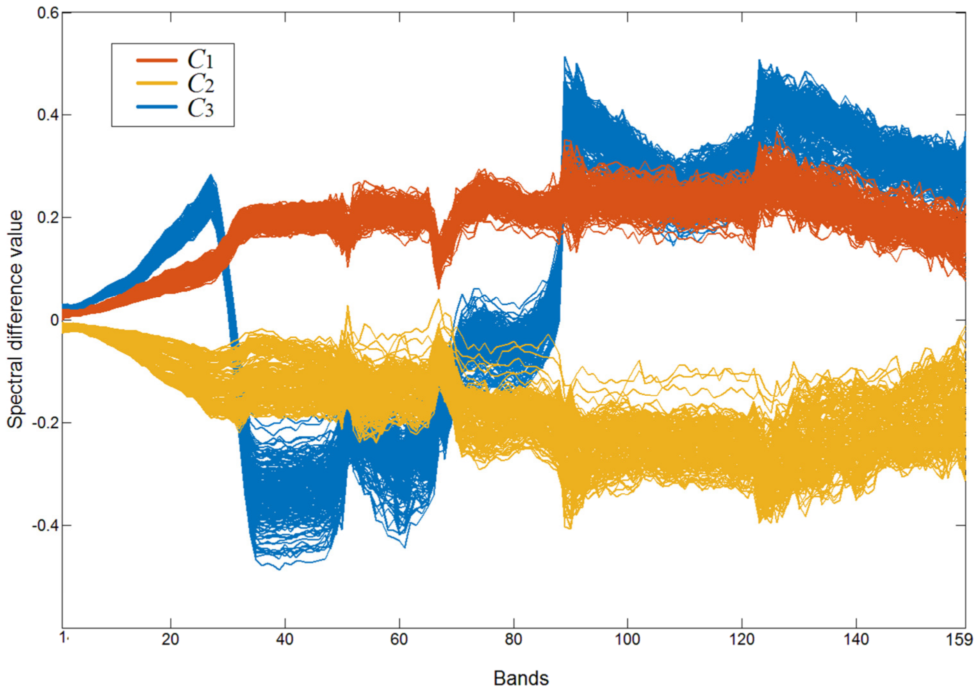
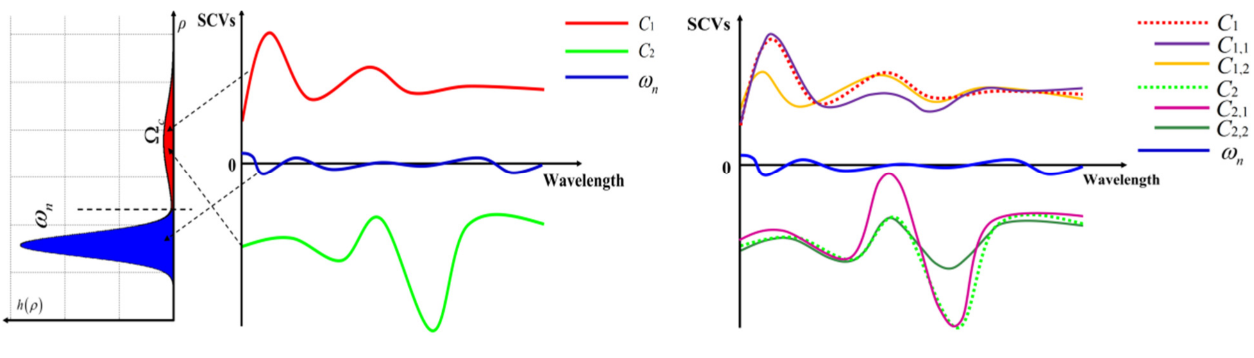


Fig.3-1 Example of the SCV signatures associated to three change classes in the spectral difference domain $\boldsymbol{X}_D$.

After a comprehensive analysis on the spectral difference domain $\boldsymbol{X}_D$, three important observations are utilized in investigating the considered multiple change detection task: 1) pixels that associated to a given change class have the same (or very similar) signatures of their SCVs thus are clustered together in $\boldsymbol{X}_D$ domain (see an example in Fig.3-1); 2) the magnitude of the SCVs drives the separation of the changed and the unchanged pixels; 3) the spectral shape of SCVs drives the discrimination of different kinds of changes. Generally speaking, a pixel can belong to the class of changed pixels in Ω_c or the one of unchanged pixels ω_n according to the magnitude of its SCV [33]. Fig.3-2 (a) gives a qualitative example of the expected behavior of the magnitude histogram of $\boldsymbol{X}_D$. Unchanged pixels show a SCV magnitude close to zero (blue mode in Fig.3-2.a). The SCV signatures of such pixels have all spectral components close to the null vector (see the blue SCV signature in Fig.3-2.b). Changed pixels show high magnitude values (red mode in Fig.3-2.a), and their SCV signatures show one or more components that are far from the null vector. It is worth noting that in the 1-Dimensional magnitude domain usually different changes contribute to a single class Ω_c, since they are highly mixed and cannot be separated according to the magnitude values (see Fig.3-2.a). A finer analysis of SCV behaviors points out that Ω_c may include contributions from several change classes (see red and green signatures associated in Fig.3-2.b) depending on how the specific kinds of changes impacted on the spectral signatures. SCVs can be preliminary separated into *major changes*. Major changes mainly depend on the land-cover class transitions and have a large spectral difference with respect to no-change class and among each other. Usually, major changes can be easily and directly identified since they significantly affect a large portion of the spectrum of hyperspectral images. In many cases they can be also detected from multispectral images. As shown in Fig.3-2 (c), each major change (i.e., C_1 and C_2) produces statistically significant different spectra compared with each other and with the class of unchanged pixels. Within each major change, it is possible to detect other clusters of pixels having significant statistical differences in some parts of the spectrum. Such clusters are defined here as s*ubtle changes*. Subtle changes have SCVs similar to a major change, but differ from it in small portions of the spectrum. In Fig.3-2 (c), subtle changes $C_{1,1}$ (in purple) and $C_{1,2}$ (in orange) belong to the same major change C_1 (in red), whereas $C_{2,1}$ (in magenta) and $C_{2,2}$ (in sea green) belong to C_2 (in green). In other words subtle changes show SCVs statistically different from each other in some components of the spectrum, but are quite similar to those of the associated major change. Subtle changes can be therefore detected only if a fine sampling of the spectral signature is available as it happens in hyperspectral images. If the sampling is poor as in the case of multispectral images, they cannot be detected.



(a) (b) (c)

Fig.3-2 Qualitative illustration of (a) the statistical distribution of the magnitude of SCVs ($h(\rho)$); the sample spectra on SCVs of (b) major changes; (c) subtle changes (solid line) within the given major changes (dotted line) that defined in the spectral difference domain $\boldsymbol{X}_D$.

According to the above discussion, $\Omega_c = \{C_1, C_2, ...\}$ is the set of major changes, i.e., changes that affect a large part of the spectrum and that have statistical properties significantly different from each other. Each major change may include subtle changes (i.e., $C_1 = \{C_{1,1}, C_{1,2}, ...\}$ and $C_2 = \{C_{2,1}, C_{2,2}, ...\}$) whereas others may not (i.e., $C_3 = \varnothing$). By iterating the process it is possible to state that each subtle change can be further split until it is not possible to detect spectrally statistical inhomogeneity. Each major or subtle change that cannot be split anymore is defined as a *change endmember*[1] in $\Omega_e=\{e_1, e_2,\ldots,e_K\}$ that associated to a given change class in $\Omega_c = \{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}\}$. Accordingly, all pixels belong to a specific change endmember have the same (or very similar) spectral behaviors in the SCV domain and thus can be clustered into the same group. Thus the problem that we need to address is related to the identification and separation of change classes in $\{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}\}$ (thus the endmembers in $\Omega_e=\{e_1, e_2,\ldots,e_K\}$) from each other and from unchanged pixels in ω_n. We assume that the considered images are all radiometric corrected, thus change endmembers are only related to the application and to the end-users. Note that the external factors (e.g., illumination conditions, seasonal effects) might have impacts on the detected change endmembers (causing differences) but almost of them will not be identified as one of those changes due to the low change magnitude.

3.2 Analysis of the Spectral Stacked Domain

Let consider a pair of bitemporal images. Under the pure spectrum assumption, pixels are spatially homogenous and thus contain only one land-cover material at each date (i.e., pixel level CD in Fig.3-3). Thus only change or no-change cases may occur according to a crisp decision strategy. However, by analyzing the bi-temporal CD problem from the perspective of mixed spectra assumption, a pixel may be associated with several possible situations of class mixtures and transitions (see Fig.3-3), thus more complex situations may occur. The spectral mixture can occur only on a single date image (i.e., $\boldsymbol{X}_1$ or $\boldsymbol{X}_2$) or on both of them, leading to the following four possible situations (see Fig.3-3, subpixel level CD): 1) the pixel is pure in both $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$; 2) the pixel is pure in $\boldsymbol{X}_1$ but mixed in $\boldsymbol{X}_2$; 3) the pixel is mixed in $\boldsymbol{X}_1$ but pure in $\boldsymbol{X}_2$; and 4) the pixel is mixed in both $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. The final crisp decision is made by assigning the change or no-change label depending on the majority of the material composition and its temporal behaviour (see Fig.3-3). If the majority is associated to different materials at the two dates, the pixel tends to be changed, whereas it tends to be unchanged when majority is associated to the same material in the two images. Thus an effective investigation at the subpixel level may point out potential spectral variations within a pixel that are usually not detectable at the pixel level. This helps to better understand the spectral

[1] Note that the definition of change endmember is in concept different from the definition of endmembers in spectral unmixing. In the latter case, endmembers are the spectral signatures of pure classes that result combined in mixed pixels due to the limited spatial resolution of the acquisition sensor.

mixture phenomenon and its effects on CD. The proposed change formulation and representation is based on this spectral mixture analysis.

The most popular change index when dealing with optical passive sensors is the B-Dimensional difference image X_D computed by subtracting the multitemporal images pixel-by-pixel [31], [86], [87], [88], (see (1) in Section 3.1). It is worth noting that the pixel spectrum in the spectral difference domain $\boldsymbol{X}_D$ has changed its physical meaning into the land-cover transitions rather than the original land-cover materials. Thus it is more complex and difficult to identify a suitable spectral mixture model (either linear or nonlinear) for $\boldsymbol{X}_D$. Moreover, in the $\boldsymbol{X}_D$ domain different kinds of no-changes might result in very similar spectral signatures (i.e., having components all close to the null vector), leading to the failure of the unmixing procedure, especially in identifying the distinct no-change endmembers. These intrinsic properties limit the effectiveness of the analysis of spectral mixture in $\boldsymbol{X}_D$. In this thesis, we change our perspective from the traditional B-Dimensional difference domain $\boldsymbol{X}_D$ (see Fig.3-4.a) into the $2B$-Dimensional multitemporal domain represented by $\boldsymbol{X}_S$ (see Fig.3-4.b), which is a stacked feature space based on the considered multitemporal images, i.e.,

$$\boldsymbol{X}_S = [\boldsymbol{X}_1, \boldsymbol{X}_2] \tag{4}$$

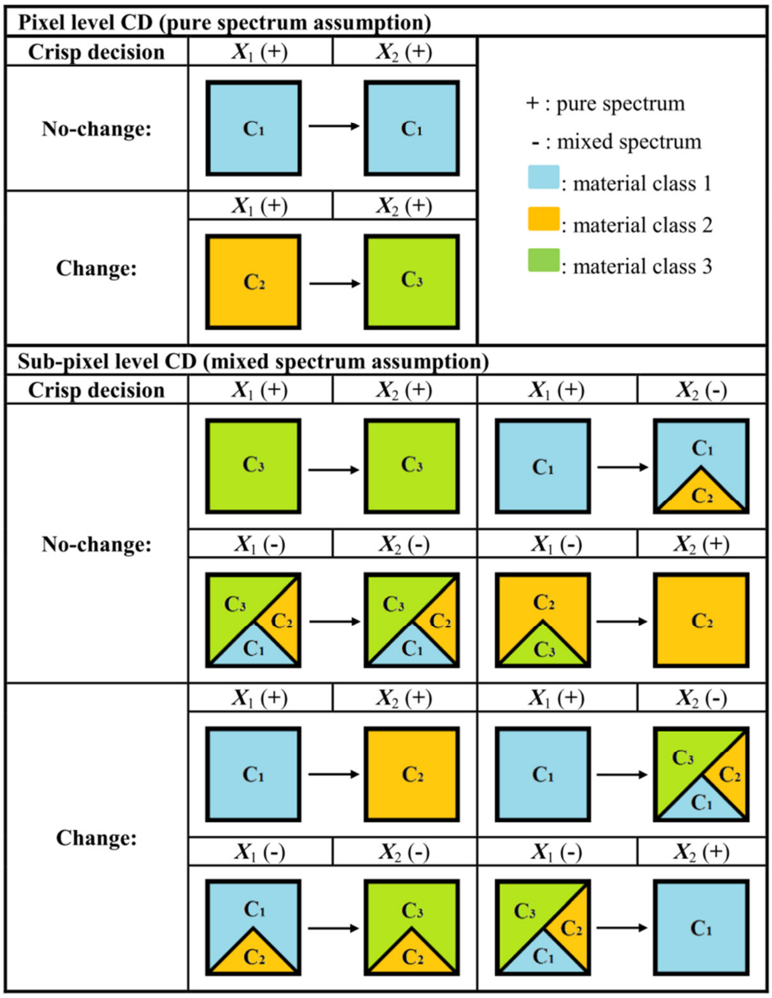


Fig.3-3 Possible change situations of a single pixel in the bi-temporal images based on the pure spectrum and mixed spectrum assumptions.

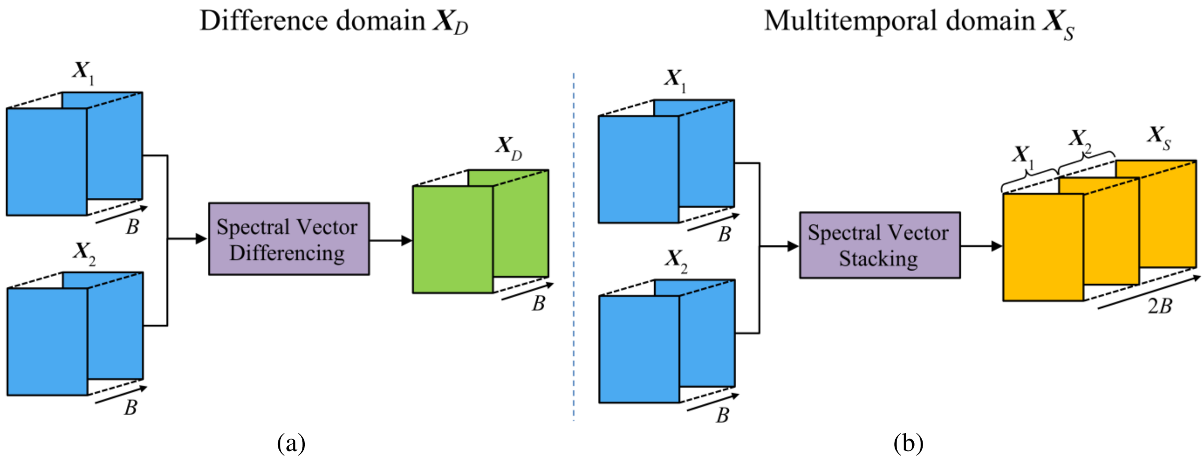


Fig.3-4 Illustration of the CD-HS problem in: (a) the B-Dimensional difference domain X_D, and (b) the $2B$-Dimensional multitemporal (stacked) domain X_S.

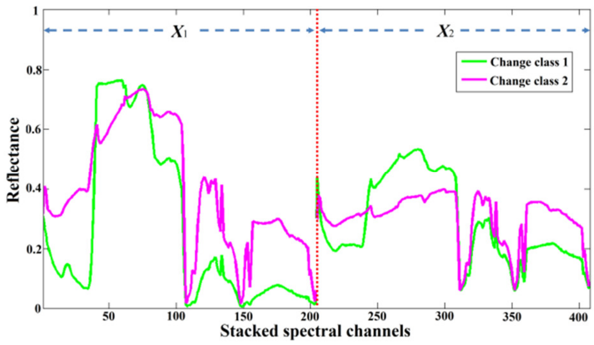
The multitemporal domain X_S has been used for CD purposes in the literature. Two main approaches can be identified: the supervised Direct Multi-date Classification (DMC) [19], [20] and the unsupervised stacked feature transformation [66], [89], [90], [91]. The former identifies changes by simultaneously classifying the stacked multi-date images, thus a land-cover transition is represented by an output class in the final classification map. Accordingly, the considered CD task is actually replaced by a supervised classification task. However, the generation of a comprehensive multitemporal training set that represents all the detailed land-cover transitions makes this supervised approach difficult to use in real applications. Sub-optimal solution to the classification problem can be implemented by using compound classification strategies [92]. The latter is implemented based on the multi-date data transformation (e.g., Temporal-Principal Components Analysis (T-PCA) [66], [89], Multi-date Kauth-Thomas (MKT) [90], Multi-date Graham-Schimidt (MGS) [91]), where a careful analysis and selection is required to find the transformed components related to the change classes of interest. Usually this step is manual and thus time consuming.

Unlike the literature works, a spectral mixture model is used in this thesis in the multitemporal domain X_S to solve the considered multiple-change detection problem. The main advantages of working in the multitemporal domain X_S are: i) it preserves the intrinsic properties of the spectral signatures that represent the real land-cover materials, which are extended along the temporal direction; and ii) only the occurred land-cover transitions are identified as endmembers in the mixture model, i.e., those that do not exist between the images are not considered. Thus a given spectral signature in X_S is defined as a mixture of the pure *multitemporal endmembers* (MT-EMs) associated to a specific kind of change or no-change class. A single spectral mixture is approximated in X_S combined from two independent mixtures in X_1 and X_2. If we consider a linear mixture model (LMM), it can be described as:

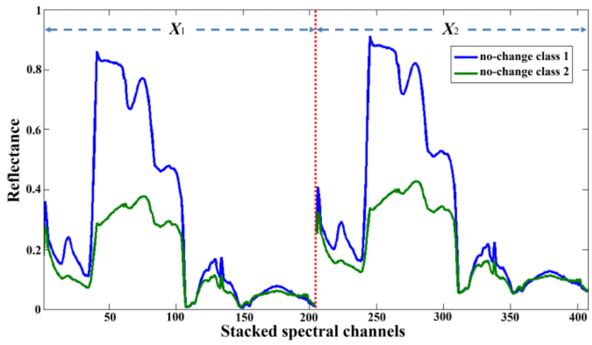
$$X_S = [X_1, X_2] = \left[\sum A_1 E_1 + N_1, \sum A_2 E_2 + N_2\right] \approx \sum A_S E_S + N_S \tag{5}$$

where E_S is the matrix of the multitemporal endmember set, A_S is the corresponding abundance matrix, and N_S represents the noise matrix. E_1 and E_2, A_1 and A_2, and N_1 and N_2 are the endmembers, abundances, and noise matrices in t_1 and t_2 image mixture models, respectively. Note that differently from the spectral signatures in $\boldsymbol{X}_D$, in the $\boldsymbol{X}_S$ domain both different change classes and different no-change classes have discriminative spectral signatures among each other. This allows one to analyze in details the contribution of different change and no-change classes to the spectral composition of the pixels.

Illustrative examples of the spectral signatures of two change classes and two no-change classes in the $\boldsymbol{X}_S$ domain are shown in Fig.3-5 (a) and (b), respectively. The spectral signatures of the change classes have different spectral shapes in the two components associated to $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ (see Fig.3-5.a), whereas two components of the spectral signatures are almost the same for the no-change classes (see Fig.3-5.b). Moreover, different change and no-change classes have distinct spectral signatures in the multitemporal domain of $\boldsymbol{X}_S$ (see Fig.3-5 a and b).



(a)



(b)

Fig.3-5 Examples of spectral signatures in the multitemporal domain $\boldsymbol{X}_S$: (a) change and (b) no-change classes.

Chapter 4

A Novel Sequential Spectral Change Vector Analysis for Discovering and Detecting Multiple Changes in Multitemporal Hyperspectral Images

This chapter presents an effective semi-automatic method for discovering and detecting multiple changes (i.e., different kinds of changes) in multitemporal hyperspectral images. Differently from the state-of-the-art techniques, the proposed method is designed to be sensitive to the small spectral variations that can be identified in hyperspectral images but usually are not detectable in multispectral images. The method is based on the proposed Sequential Spectral Change Vector Analysis (S^2CVA), which exploits an iterative hierarchical scheme that at each iteration discovers and identifies a subset of changes. The proposed approach is developed in an interactive and semi-automatic fashion that allows one to investigate in detail the structure of changes hidden in the variations of the spectral signatures according to a systematic top-down procedure. The proposed approach has been tested on three hyperspectral data sets including both simulated and real multitemporal images showing multiple-change detection problems. Experimental results confirmed the effectiveness of the proposed method.

4.1 Introduction

For decades, Earth Observation satellites have provided a unique way to observe our living planet from space. Thanks to the revisit property of EO satellites, a huge amount of multitemporal images is now available in archives. This allows us to monitor the land surface changes in wide geographical areas according to both long term (e.g., yearly) and short term (e.g., daily) observations. The detection and understanding of changes occurred in the multitemporal images is essential for studying the global change, the environmental evolution and the anthropic phenomena [6].

For the unsupervised multiple-change detection techniques, Transformation-based techniques like Iteratively Reweighted Multivariate Alteration Detection (IR-MAD) [43], Temporal-Principal Components Analysis (T-PCA) [66], etc., were proposed and proven to be effective in the literature. We recall the Compressed Change Vector Analysis (C^2VA) approach recently presented in the literature [47], which was developed based on the polar Change Vector Analysis (CVA) [33], [44]. Unlike CVA, the C^2VA allows a visualization and detection of multiple changes by considering all the available spectral channels within a 2-Dimensional representation. Thus C^2VA method theoretically allows one to detect all possible change classes occurred between the considered images, without neglecting any spectral band or working

This chapter is accepted for publication for *IEEE Transactions on Geoscience and Remote Sensing.* Title : “Sequential Spectral Change Vector Analysis for Iteratively Discovering and Detecting Multiple Changes in Hyperspectral Images”. Authors: S.Liu, L.Bruzzone, F.Bovolo, M.Zanetti, P.Du.

on their selection. Despite the C^2VA representation exploits a lossy compression of the information, it has proven to be successful in addressing multiple-change detection problems in multispectral images [47], [48].

The growing availability of hyperspectral data brings the remote sensing into a high spectral resolution era. Hyperspectral sensors take images having a very high spectral resolution (e.g., 10 nm) over a wide spectral range (e.g., 400 nm-2500 nm). In change detection, this important property allows one to potentially detect small spectral variations that are usually not detectable in multispectral images due to the poor spectral representation (i.e., generally sufficient for representing only the major abrupt changes) [93]. Accordingly, robust CD techniques should be developed to take full advantage of the rich spectral information contained in hyperspectral data, and to effectively identify the multiple-change information. In this chapter we focus on the problem of representation and analysis of multiple-change information in hyperspectral images.

Despite the successful definition of a large number of effective CD techniques for multispectral images, these techniques reduce their efficiency when hyperspectral images are considered mainly due to: 1) the high-dimensionality of the feature space; 2) the presence of noisy channels and redundant information; 3) the increase of computational cost; 4) the increase of the possible number of changes; and 5) the high complexity of the change representation and identification process. In particular, the last two items may strongly affect the effectiveness of CD methods, like the C^2VA, because it might be highly difficult (in some cases impossible) to: i) identify successfully all the existing change clusters and thus the correct number of changes; and ii) model and extract each single change class. Therefore, more advanced and sophisticated approaches should be designed to properly handle the challenging issues in multitemporal hyperspectral images. Recently, we proposed an unsupervised hierarchical spectral change vector analysis (HSCVA) method that addresses the considered problem via a hierarchical clustering procedure [86]. At each level of the processing, automatic clustering is applied based on the principal components to estimate the number of changes and to discriminate them in different clusters. Although the usefulness of HSCVA has been proven in real CD-HS cases, HSCVA does not allow an explicit, detailed and interactive analysis of the changes in the spectral signatures of multitemporal images.

In this chapter, the multiple-change detection problem in hyperspectral images is analyzed from the spectral signature point of view. The limitations that result in the degradation of performance when C^2VA is applied to hyperspectral images are studied. A novel Sequential Spectral Change Vector Analysis (S^2CVA) approach is proposed, which: 1) discovers and analyzes the multiple changes at different spectral levels through a top-down hierarchical architecture; 2) at each level provides a visualization of multiple changes in a 2-Dimensional representation domain; 3) is designed in a sequential, interactive and semi-automatic fashion. In detail, the proposed approach addresses the CD problem as follows. First, a binary CD step is applied to hyperspectral multitemporal images to extract in a conservative way the changed pixels from the whole difference image. Then the attention is focused only on the changed pixels, and a novel adaptive 2-Dimensional change representation is proposed to visualize the possible different changes. Changes are separated between each other according to an interactive change identification scheme. Then the process moves to the next level and iterates on each of the identified changes until convergence is reached. Finally, the change-detection map is generated by merging the detected change classes derived at each level of the hierarchy. Note that at each step of the hierarchical processing, the proposed technique emphasizes a specific portion of the change information in the whole Spectral Change Vector (SCV) feature space. This is accomplished by adaptively generating proper change variables for the 2-Dimensional change representation. The proposed S^2CVA approach is validated on three data sets including: 1) simulated bi-temporal images based on an AVIRIS hyperspectral image; 2) real bi-temporal hyperspectral images acquired by the Hyperion sensor onboard of EO-1 satellite; and 3) simulated bi-temporal images based on a hyperspectral camera image. Experimental results confirm the effectiveness of the proposed method for addressing the multiple-change detection problem in multitemporal hyperspectral images.

The rest of the chapter is organized as follows. Section 4.2 reviews the C^2VA approach and discusses some important issues and challenges when transferring the CD perspective from multispectral to hyperspectral cases. The proposed sequential CD technique (i.e., S^2CVA) is described in Section 4.3. Section 4.4 introduces the used hyperspectral data sets. Section 4.5 reports and analyzes the obtained experimental results. Finally, Section 4.6 draws the conclusion of this chapter.

4.2 Multiple-Change Detection by C^2VA in Multi/Hyper-spectral Images

4.2.1 Standard C^2VA

In the standard C^2VA method, a compressed change representation in a 2-Dimensional polar domain is defined by two change variables, i.e. the magnitude ρ and the direction α [47] :

$$\rho=\sqrt{\sum_{b=1}^{B}\left(X_D^b\right)^2} \tag{6}$$

$$\alpha=\arccos\left[\frac{1}{\sqrt{B}}\left(\sum_{b=1}^{B}X_D^b \middle/ \sqrt{\sum_{b=1}^{B}\left(X_D^b\right)^2}\right)\right] \tag{7}$$

where X_D^b is the b-th (b=1,…,B) component of X_D and B is the number of spectral channels of the considered images (i.e., the dimensionality of SCVs). The magnitude ρ is defined based on the popular Euclidean distance [93]. It measures the total contribution of spectral change brightness, whereas it is not sensitive to the shape of spectral vectors. The angle distance α is measured by the Spectral Angle Distance (SAD) [94], which is widely used in several hyperspectral application fields for material identification, classification, etc. [94], [95], [96]. SAD measures the similarity between two given spectral signatures, especially focusing on the shape of the spectrum. More details about the two distance metrics are given in Appendix A.

Variables ρ and α define a 2-Dimensional polar coordinate domain $\boldsymbol{D}$ [47] as:

$$\boldsymbol{D} = \{\rho \in [0, \rho_{\max}] \text{ and } \alpha \in [0, \pi]\} \tag{8}$$

where ρ_{max} is the maximum value of ρ. All the SCVs in $\boldsymbol{X}_D$ can be represented in a 2-Dimensional semi-circle scattergram (see Fig.4-1). This scattergram allows one to easily visualize multidimensional change information in a 2-Dimensional feature space. However, the compression from a B-Dimensional space into a 2-Dimensional space results in a loss of information and thus in ambiguity on the detection of different kinds of changes.

In the C^{2}VA framework, the multiple-change detection problem is addressed according to two steps [47]:

1) Set threshold T_ρ along ρ variable to divide the whole semicircle of the C^{2}VA representation domain into two parts, i.e., SC_n and SA_c (see Fig.4-1), which are related to the unchanged (ω_n) and changed (Ω_c) SCVs, respectively:

$$SC_n = \{\rho, \alpha | 0 \le \rho < T_\rho \text{ and } 0 \le \alpha \le \pi\} \tag{9}$$

$$SA_c = \{\rho, \alpha | T_\rho \le \rho \le \rho_{\max} \text{ and } 0 \le \alpha \le \pi\} \tag{10}$$

2) Separate multiple-change classes ($\omega_{c_1}, ..., \omega_{c_K}$) along the direction α by analyzing the semiannulus SA_c. Multiple angular thresholds $T_{\alpha,k}$ (k=1,…,K-1) can be defined to find K annular sectors (each corresponding to a change) inside SA_c:

$$S_k = \{\rho, \alpha | T_\rho \le \rho < \rho_{\max} \text{ and } T_{\alpha,k} \le \alpha \le T_{\alpha,k+1}\} \tag{11}$$

where $0 \le T_{\alpha,k} < T_{\alpha,k+1} \le \pi$.

Note that thresholds T_ρ and T_α can be detected manually or automatically according to one of the various methods proposed in the literature [31], [92], [97], [35].

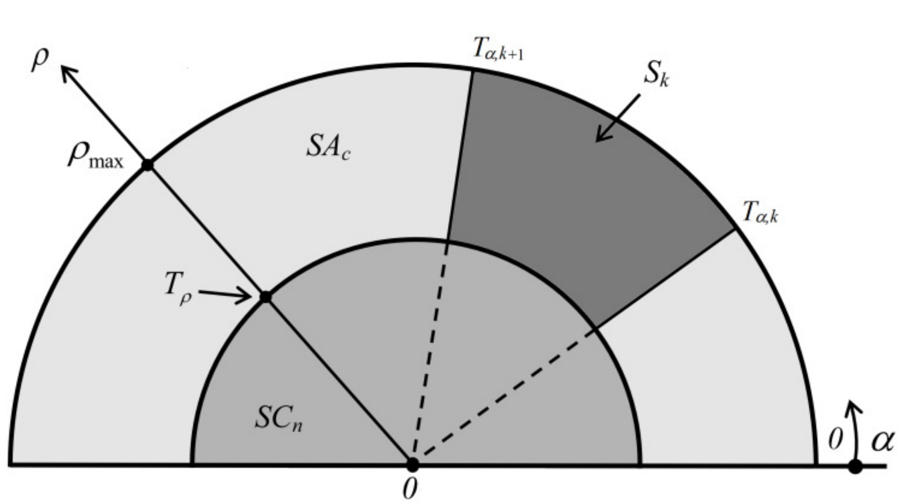


Fig.4-1 Compressed Change Vector Analysis (C^{2}VA) for representing multiple-changes in the 2-D polar domain [47].

4.2.2 Problems and Challenges When Applying C^2VA to CD-HS Cases

C^2VA has proven to be effective when dealing with the multispectral images [6], [47] , [48]. Due to the rough spectral resolution of multispectral images, the direction variable α is generally effective to represent the relatively few abrupt changes that are visible in multispectral data and does not suffer too much of the loss of information associated with the compressed representation. However, in the CD-HS case the finer spectral resolution results in a high number of changes that can be detected [86]. Thus the compressed representation of many spectral channels in one variable α leads to a high probability of ambiguous description of changes, which may result in highly overlapped clusters in the C^2VA representation.

In multitemporal hyperspectral images, as we defined in Chapter 3, we can distinguish between two kinds of hierarchically related changes: 1) major changes, which have a significant spectral difference with respect to both the no-change class and the other change classes; 2) subtle changes, whose SCVs are similar to those of an associated major change, but statistically significantly differ from each other in some portions of the spectrum. When subtle changes inside a major change are present, due to a high similarity among their SCVs, the α variable defined on the basis of a fixed unit reference vector $\boldsymbol{R}$ = $[1/\sqrt{B}$ $,\ldots, 1/\sqrt{B}]$ (as the one derived in (7), [47]) is likely to be ineffective for their discrimination. To overcome this drawback, a robust definition to the reference vector $\boldsymbol{R}$ (and thus the change variable α) should be designed. This definition should optimize the separation of changes and provide a meaningful change visualization in a hierarchical and adaptive way, thus to properly represent and discover as many changes present in the hyperspectral images as possible.

Accordingly, let use recall that the angle distance α between SCVs in $\boldsymbol{X_D}$ and a generic reference vector $\boldsymbol{R}$ is defined as:

$$\alpha = \arccos\left[\left(\sum_{b=1}^{B}\left(X_D^b R^b\right) \Big/ \sqrt{\sum_{b=1}^{B}\left(X_D^b\right)^2 \sum_{b=1}^{B}\left(R^b\right)^2}\right)\right], \qquad \alpha \in [0,\pi] \tag{12}$$

where R^b is the b-th component of the reference vector $\boldsymbol{R}$. Due to the fact that α is invariant to the multiplicative scaling, the reference $\boldsymbol{R}$ in [47] actually can be extended to any constant vector $\boldsymbol{R}$ = $[a,\ldots,a]$, where $a > 0$ and $a \in \Re$. More details can be found in Appendix A.

4.3 Proposed Approach

Inspired by the aforementioned analysis and discussion on C^2VA, we are motivated to find a reliable solution to the problem of discovering, representing and discriminating different spectral changes in hyperspectral images. The idea is to start from the C^2VA developed for multispectral images, and to overcome its drawbacks when applied to hyperspectral images. To this end, a novel Sequential Spectral Change Vector Analysis (S^2CVA) is proposed. Differently from the standard C^2VA, the proposed S^2CVA is designed to be sensitive to the small changes in spectral signature behaviors, which are usually not detectable in multispectral images. Instead of having just a one shot processing as in the standard C^2VA (due to the measurement based on the fixed reference vector), the proposed S^2CVA allows an adaptive definition of the reference vector. The whole S^2CVA framework is designed in an iterative fashion. A novel unsupervised 2-Dimensional Adaptive Spectral Change Vector Representation (ASCVR) technique and a fast change identification scheme are designed aiming at better representing and discovering the possible multiple changes at each analysis level. This results in a sequence of 2-Dimensional representation scattergrams that can be obtained to model the multiple changes at different levels of the hierarchy taking into account the global and local spectral variations. Homogenous scattering clusters observed from a change representation scattergram at a given level are discriminated among each other and then are investigated individually in the next level. Note that the proposed S^2CVA is developed as a semi-automatic technique, which consists of: i) an adaptive change representation that permits to discover the multiple changes, and ii) a manual (interactive) change identification. The block scheme of the proposed CD approach is illustrated in Fig.4-2.

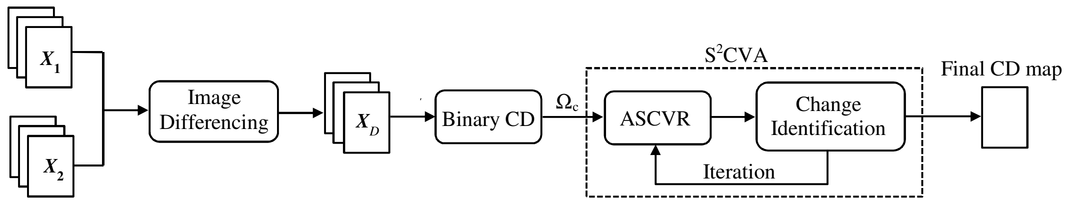


Fig.4-2 Block scheme of the proposed CD approach based on S^2CVA.

At the beginning, all SCVs in $\boldsymbol{X}_D$ belong to the same class Ω. Once the binary CD result is obtained in level L_0, let P_0 be the cluster representing SCVs associated with the general change cluster Ω_c extracted to initialize the whole sequential analysis. Note that in L_0, the binary CD is done as in the standard C^2VA method, where only the magnitude variable ρ is analyzed according to a proper thresholding technique (e.g., by using the EM algorithm in the framework of the Bayesian decision theory [31], [47]). In the next levels, let $P_{h,j}$ be the j-th (j = 1,…,J_h) change cluster observed at the h-th (h=1,…,H) level L_h of the S^2CVA hierarchy, where J_h is the maximum number of change clusters at level h, and H is the total number of levels in the hierarchy. In the next subsections, firstly we define the ASCVR technique and then describe its use in the framework of the hierarchical analysis associated with the proposed S^2CVA.

4.3.1 Proposed 2-D Adaptive Spectral Change Vector Representation (ASCVR)

Similarly to C^2VA, the proposed ASCVR technique is designed in a 2-Dimensional feature space. The main reason for using a 2-Dimensional rather than a high dimensional (e.g., B-Dimensional) feature space is that in this way it is easy to visualize the change clusters and their numbers. Instead of using a fixed reference vector $\boldsymbol{R}$ as in C^2VA [47], the proposed ASCVR is designed to adaptively define the most suitable reference vector $\boldsymbol{R}$ for analyzing the multiple changes at each considered hierarchical level. The selection of $\boldsymbol{R}$ and the change representation are directly derived based on the statistic distribution of the input SCVs. Thus the proposed ASCVR is totally unsupervised.

The proposed ASCVR is defined by the change magnitude ρ (which is the same as defined in C^2VA (6)) and the compressed change direction α (which represents the spectral angle distance between a given SCV in $\boldsymbol{X}_D$ and a reference vector $\boldsymbol{R}$ (12)). Different values of angle distance measure indicate different kinds of changes. At different levels of the proposed hierarchical analysis, for each specific portion of the SCV feature space associated with the considered change cluster, a new reference vector $\boldsymbol{R}$ is defined. In greater details, the first eigenvector that corresponds to the maximum eigenvalue of the data covariance matrix of the SCVs associated to the considered cluster j at level h (i.e., $P_{h,j}$) is adopted as the reference vector. The selection of the first eigenvector is due to the fact that a 2-Dimensional change representation is desired, which preserves as much as possible the spectral variations of the considered SCVs in a low-dimensional feature space that can be easily managed. In the proposed ASCVR, this vector shows a direction that maximizes the variance of the measurement on α, thus resulting in an adaptive and effective representation of the hidden change patterns in the considered SCVs. For a considered generic cluster $P_{h,j}$ in the hierarchy, the procedure for adaptively defining the reference vector $\boldsymbol{R}_{h,j}$ is as follows. Let us consider the covariance matrix $\boldsymbol{\Gamma}_{h,j}$ of $\boldsymbol{x}_{h,j}$ (denoted as the SCVs in $P_{h,j}$):

$$\boldsymbol{\Gamma}_{h,j} = \mathrm{cov}(\boldsymbol{x}_{h,j}) = E\left[(\boldsymbol{x}_{h,j} - E[\boldsymbol{x}_{h,j}])(\boldsymbol{x}_{h,j} - E[\boldsymbol{x}_{h,j}])^T \right] \tag{13}$$

where $E[\boldsymbol{x}_{h,j}]$ is the expectation of $\boldsymbol{x}_{h,j}$ and $\boldsymbol{\Gamma}_{h,j}$ is the $B \times B$ dimensional covariance matrix that represents $\boldsymbol{x}_{h,j}$ by means of the eigenvectors and eigenvalues, which are calculated according to (14):

$$\boldsymbol{\Gamma}_{h,j} \cdot \boldsymbol{V}_{h,j} = \boldsymbol{V}_{h,j} \cdot \boldsymbol{W}_{h,j} \tag{14}$$

$\boldsymbol{W}_{h,j}$ is a diagonal matrix where the eigenvalues are sorted in descending order (i.e., $\lambda_{h,j}^1 > \lambda_{h,2}^2 > \ldots > \lambda_{h,j}^B$) in the diagonal. The magnitude of eigenvalues reflects the amount of data variance that is captured by the corresponding eigenvectors. Let $\boldsymbol{V}_{h,j} = \left[\boldsymbol{V}_{h,j}^1, \boldsymbol{V}_{h,j}^2, \ldots, \boldsymbol{V}_{h,j}^B \right]$ be the matrix of eigenvectors. The reference vector $\boldsymbol{R}_{h,j}$ for computing $\alpha_{h,j}$ in (12) is selected as the first eigenvector $\boldsymbol{V}_{h,j}^1$ that corresponds to the largest eigenvalue $\lambda_{h,j}^1$:

$$\boldsymbol{R}_{h,j} = \boldsymbol{V}_{h,j}^1 = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_B \end{pmatrix}_{h,j} \tag{15}$$

Accordingly, the two change variables (i.e., magnitude $\rho_{h,j}$, direction $\alpha_{h,j}$) for the SCVs of the cluster $P_{h,j}$ are defined as follows:

$$\begin{cases} \rho_{h,j} = \sqrt{\sum_{b=1}^{B} \left(\boldsymbol{x}_{h,j}^{b} \right)^2} \\ \alpha_{h,j} = \arccos\left[\left(\sum_{b=1}^{B} \left(\boldsymbol{x}_{h,j}^{b} R_{h,j}^{b} \right) \middle/ \sqrt{\sum_{b=1}^{B} \left(\boldsymbol{x}_{h,j}^{b} \right)^2 \sum_{b=1}^{B} \left(R_{h,j}^{b} \right)^2} \right) \right] \end{cases} \tag{16}$$

The resulted 2-Dimensional representation domain $\boldsymbol{D}_{h,j}$ is defined as:

$$\boldsymbol{D}_{h,j} = \left\{ \rho_{h,j} \in [0, \rho_{h,j}^{\max}] \; and \; \alpha_{h,j} \in [0, \pi] \right\} \tag{17}$$

where $\rho_{h,j}^{\max}$ is the maximum change magnitude value of $\rho_{h,j}$.

Note that the use of the first eigenvector does not guarantee the maximum discriminability in all cases, but this is an effective choice for an adaptive 2-D visualization of the latent change information. An alternative way that mathematically approximates to the use of first eigenvector is by using the Singular Value Decomposition (SVD) [87]. SVD is an orthogonal linear transformation that is able to capture the underlying variance of the data. This important property makes SVD effective in wide applications like data compression, feature extraction, etc., especially it is very useful for dimensionality reduction of the high-dimensional data and providing meaningful data visualization [98]. More details for the extracting of the reference vector $\boldsymbol{R}$ by SVD are given in the Appendix B.

4.3.2 Proposed Sequential Spectral Change Vector Analysis (S^{2}CVA)

The intrinsic adaptive characteristic of the proposed ASCVR can be exploited to represent either the whole or a portion of the SCV space. Accordingly, the challenging CD-HS problem can be addressed hierarchically by the proposed Sequential Spectral Change Vector Analysis (S^{2}CVA). The main idea is that a sequence of ASCVR scattergrams can be obtained for representing SCVs in different specific spectral levels, thus SCVs that are associated with different change classes can be gradually separated following the sequential analysis. The block scheme of the proposed S^{2}CVA method is shown in Fig.4-3.

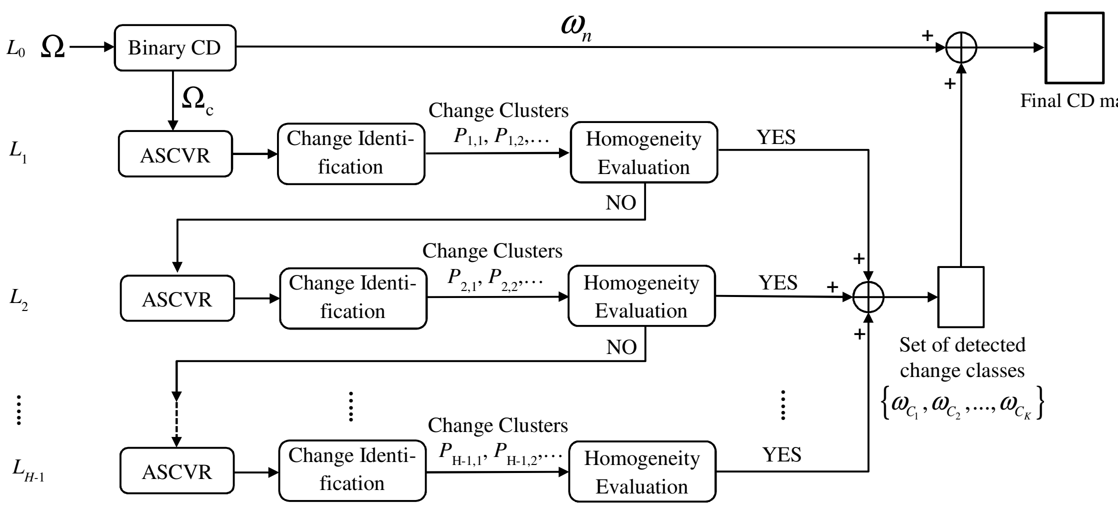


Fig.4-3 Block scheme of the proposed S^{2}CVA step.

The process begins at the initial level L_0 of the hierarchy, where P_0 is the initial change cluster obtained by thresholding the ρ variable. Then the sequential technique focuses on the P_0 (i.e., SCVs that belong to Ω_c) in order to discover and identify the possible kinds of changes. At level L_1, if more than one homogenous clusters $P_{1,1}$, $P_{1,2}$,…, P_{1,J_1} is observed, changes need to be discriminated and separated among each other. Then the procedure moves to the next level L_2, focusing on each single cluster in {$P_{1,1}$, $P_{1,2}$,…, P_{1,J_1} } to continue the iterative change analysis by re-projecting the SCVs of each cluster in its corresponding new ASCVR domain. Thus this is to explore if more possible changes at this level of representation can be discovered. The ASCVR scheme defined in the Section 4.3.1 is used for each considered change cluster, thus the corresponding 2-Dimensional change representations $D_{1,j}$ are built by automatically updating the references $\boldsymbol{R}_{1,j}$ (and thus $\alpha_{1,j}$) according to the SCVs $\boldsymbol{x}_{1,j}$ in $P_{1,j}$ (j=1,…,J_1). By considering the intrinsic properties of SCVs that SCVs belong to a given change class have homogenous behaviors on the change variables, thus a single cluster are expected to be observed on their representation domain. On the contrary, different clusters that can be discriminated in the ASCVR scattergram indicate possible different kinds of changes. A manual procedure is applied interactively by the user. For each observed change cluster, a discrimination boundary is manually defined, which is structured as a polygon in the software prototype we implemented. Note that boundaries are selected independently for each observed change cluster. At a given level of S^2CVA the homogeneity is evaluated manually. The convergence is reached when only a single homogenous change cluster is observed in the scattergram. The cluster is finally associated to a change class in $\left\{\omega_{c_1},\omega_{c_2},...,\omega_{c_K}\right\}$ and the SCVs of that change class are reversely mapped into the image space to generate the CD map. The whole CD procedure is completed when each representation of the considered specific portion of SCVs in the hierarchy achieves convergence. The whole S^2CVA hierarchy can be modeled as a tree-structure (see the example in Fig.4-4). The final multiple-change detection map is the union of all the detected change classes $\left\{\omega_{c_1},\omega_{c_2},...,\omega_{c_K}\right\}$ and the extracted no-change class ω_n (i.e., the union of all the leaf nodes in the tree). It is worth noting that by following the proposed sequential analysis, the original global change representation and optimization problem is decomposed into several local sub-problems at different levels. Thus many potential changes can be detected hierarchically taking into account different levels of spectral change significance.

The proposed S^2CVA method allows the user to have effective interactions with the change representation and discovery (in the 2-Dimensional representation domain), and the change extraction (in the original image domain). At the end of the process the obtained hierarchical tree completely describes major and subtle changes present in the considered hyperspectral multitemporal images and their parental relations.

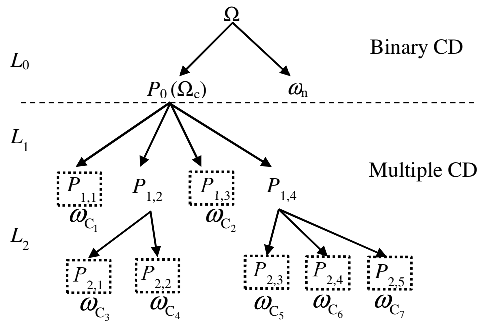


Fig.4-4 Example of the obtained three-level hierarchical tree by the proposed CD method based on S[2]CVA, where seven of the leaf nodes are detected as change classes and one as the no-change class.

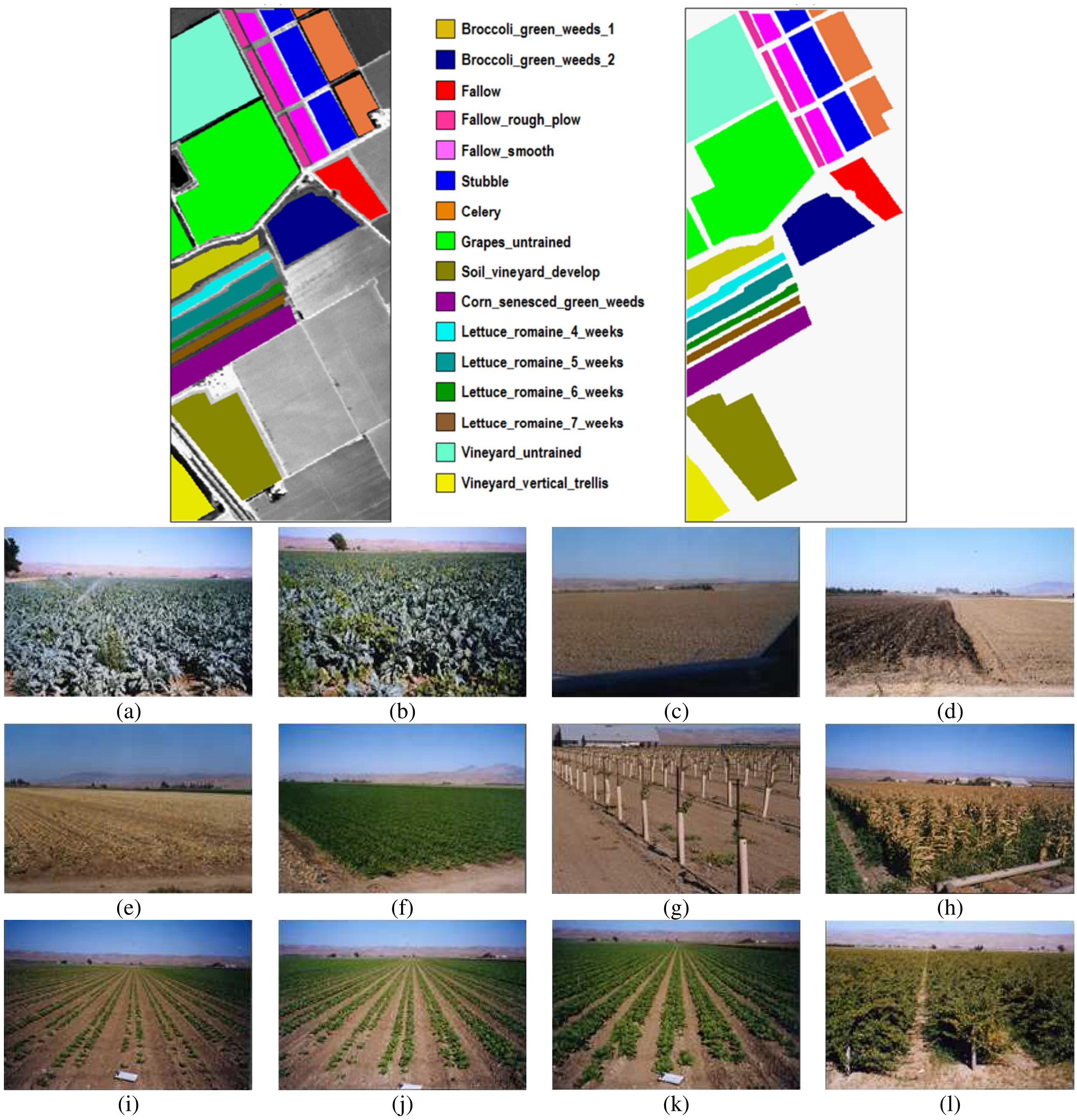


Fig.4-5 The AVIRIS Salinas data set and the reference ground truth data. Row 1: The land-cover classes in the Salinas scene. Row 2: Photographs taken at the site during data collection [99], including (a) Brocoli_green_weeds_1; (b) Brocoli_green_weeds_2; (c) Fallow; (d) Fallow_rough_smooth; (e) Stubble; (f) Celery; (g) Soil_vineyard_develop; (h) Corn_senesced_green_weeds; (i) Lettuce_romaine_5_weeks; (j) Lettuce_romaine_6_weeks; (k) Lettuce_romaine_7_weeks; (l) Vineyard_untrained.

4.4 Data Set Description and Design of Experiments

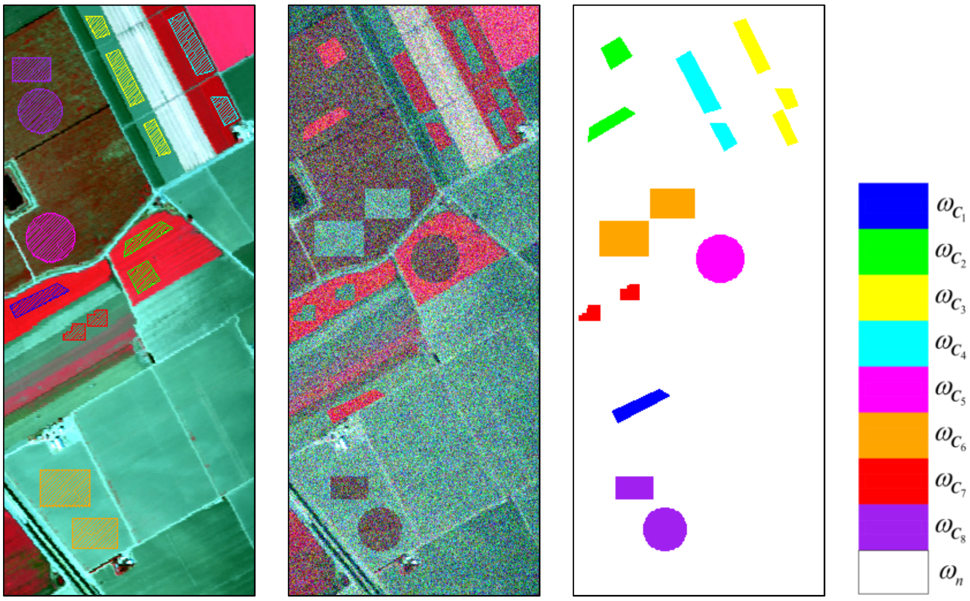
4.4.1 Description of Data Sets

Data Set 1: Simulated Hyperspectral Remote Sensing Data Set

The first data set is made up of a real single-time hyperspectral image acquired by the AVIRIS sensor in 1998 over Salinas Valley, California. This image was downloaded from the website of computational intelligence group from the Basque University (UPV/EHU) [100]. The original image contains 224 contiguous spectral bands with wavelength from 400 nm to 2500 nm. The image is characterized by a spatial resolution of 3.7 m and a spectral resolution of 10 nm and has a size of 512×217 pixels. This data set was originally used for testing a hyperspectral image classification task with the available ground truth that has 16 classes mainly including vegetation, bare soil, and vineyard (see Fig.4-5). In the pre-processing phase, 20 water absorption bands (i.e., bands 108-112, 154-167 and 224) were discarded thus obtaining 204 bands for our experiments. By taking advantage of the available ground truth data, we simulated and generated the changed image (considered as $\boldsymbol{X}_2$) based on the original image (considered as $\boldsymbol{X}_1$) according to the following steps: i) 15 tiles (i.e., regions) were extracted from the original image $\boldsymbol{X}_1$ (see Fig.4-6.a), which cover different land-cover classes. ii) The extracted tiles were inserted in different areas on $\boldsymbol{X}_1$ by replacing the spectral vectors over all bands. The same operation was done for all tiles to simulate an image ($\boldsymbol{X}_2$) with eight different change classes. iii) A small constant bias value was applied to $\boldsymbol{X}_2$ to simulate a stationary difference in light condition. iv) White Gaussian noise was added to $\boldsymbol{X}_2$ by setting an SNR equal to 10 dB. The reason for testing with the simulated data is that all the details can be quantitatively investigated in a controlled environment. False color composites of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are shown in Fig.4-6 (a) and (b), respectively. The reference map is reported in Fig.4-6 (c). Detailed class transitions are listed in TABLE 4-1 with the corresponding number of samples in each simulated change class.

TABLE 4-1 SIMULATED CHANGES IN SALINAS DATA SET AND THE CORRESPONDING NUMBER OF SAMPLES

Change class	Simulated changes (from X_1 to X_2)	Samples (Number of pixels)
ω_{C_1}	Brocoli_green_weeds_1 → Corn_senesced_green_weeds	534
ω_{C_2}	Brocoli_green_weeds_2 → Vinyard_untrained	878
ω_{C_3}	Fallow_smooth → Celery	1149
ω_{C_4}	Celery → Fallow_smooth	1163
ω_{C_5}	Grapes_untrained → Brocoli_green_weeds_2	1402
ω_{C_6}	Soil_vinyard_develop → Grapes_untrained	2347
ω_{C_7}	Lettuce_romaine_5wk → Brocoli_green_weeds_1	420
ω_{C_8}	Vinyard_untrained → Soil_vinyard_develop	1812
ω_n	No-change	101399



(a) (b) (c)

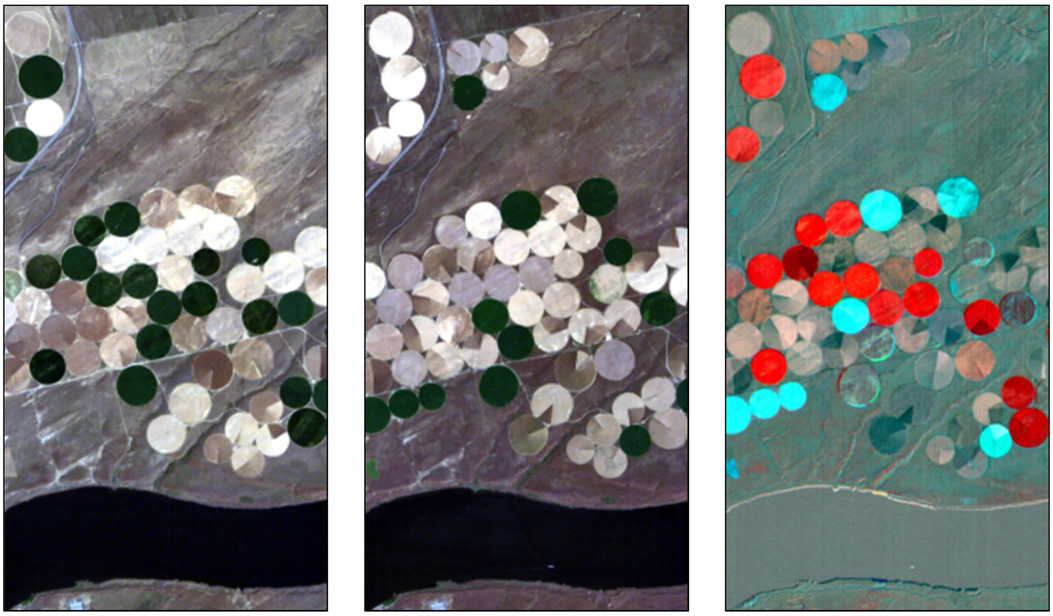
Fig.4-6 False color composite (Bands: R: 40, G: 30, B: 20) of (a) the real hyperspectral image acquired by the AVIRIS sensor in Salinas scenario ($\boldsymbol{X}_1$) and (b) the simulated changed image ($\boldsymbol{X}_2$) computed with an additive white Gaussian noise; (c) the change reference map (eight changes in different colors, and no-change class in white color).

(a) (b) (c)

Fig.4-7 False color composite (R: 710 nm; G: 620 nm; B: 510 nm) of (a) the hyperspectral image acquired by the Nuance FX hyperspectral camera (X_1) and (b) simulated image with changes (X_2). (c) Change reference map (ten changes in different colors, no-change class in white color).

Data Set 2: Simulated Hyperspectral Camera Data Set

The second data set is taken from a real-world database of hyperspectral images, which includes images acquired by a commercial hyperspectral camera (Nuance FX, CRI Inc.) [101]. With an integrated liquid crystal tunable filter, the camera acquires hyperspectral images by sequentially tuning the filter through a series of 31 narrow wavelength bands. The bandwidth is approximately 10nm in a wavelength range from 420nm to 720nm, covering mainly the visible spectrum region. The selected image is an outdoor scene in the Harvard University with a size of 1392×1040 pixels (see Fig.4-7.a). Based on the original image ($\boldsymbol{X}_1$), eight tiles were extracted over all the spectral bands and inserted into disjoint areas on a copy of $\boldsymbol{X}_1$. Thus a synthetic image ($\boldsymbol{X}_2$) was generated, which includes ten change classes. A small constant bias value was applied to $\boldsymbol{X}_2$ and white Gaussian noise was added to $\boldsymbol{X}_2$ with an SNR value equal to 20dB. The false color composite of images $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are shown in Fig.4-7 (a) and (b), respectively. Fig.4-7 (c) presents the change reference map. Note that changes were simulated considering either a material transitions or different illumination conditions for the same material. Thus subtle changes were introduced, and the complexity of the considered problem increased.



(a) (b) (c)

Fig.4-8 Real bi-temporal Hyperion images acquired on an agricultural scenario. False color composite (wavelength: R: 650.67nm, G: 548.92nm, B: 447.17nm) of the images acquired in (a) 2004 ($\boldsymbol{X}_1$) and (b) 2007 ($\boldsymbol{X}_2$). (c) Composite of three SCV channels (R: 1729.70nm, G: 1023.40nm, B: 752.43nm).

Data Set 3: Real Hyperion Remote Sensing Satellite Data Set

The third data set is made up of a pair of real bi-temporal hyperspectral remote sensing images acquired by the Hyperion sensor mounted onboard the EO-1 satellite on May 1, 2004 ($\boldsymbol{X}_1$) and May 8, 2007 ($\boldsymbol{X}_2$), respectively. Images were downloaded from the U.S. Geological Survey (USGS) website [102]. The study area is an agricultural land of Hermiston city in Umatilla County, Oregon, United States. The selected area, which has a size of 211×396 pixels, is a subset of the original whole image. The original images contain 242 spectral channels, whose wavelength range is from 350nm to 2580nm. The images are characterized by a spectral resolution of 10nm and a spatial resolution of 30m. After the pre-processing phase (e.g., uncalibrated and noisiest bands removal, bad stripes repairing, atmospheric corrections, co-registration, etc.), 159 pre-processed bands (i.e., bands: 8-57, 82-119, 131-164, 182-184, 187-220) were used for testing the proposed CD approach. For more details on the data set and the pre-processing operations readers are referred to [86]. Fig.4-8 (a) and (b) show a false color composite of the two images. The changes occurred in the considered images include land-cover transitions between crops, bare soil, water, variations in soil moisture and in the water content of vegetation. For example, the circular fields (see Fig.4-8) change their spectral signatures mainly due to the effect of the agricultural irrigation system. In this case no ground truth data are available. Thus validation of the results was done in a qualitative way by a detailed visual comparison. Fig.4-8 (c) presents a false color composition of $\boldsymbol{X}_D$ by using three selected channels. Different colors represent the possible kinds of changes, whereas the gray areas indicate the unchanged pixels. In this work we are interested in detecting all kinds of changes affecting the spectral signatures. Note that he false color composition only shows the presence of changes in the considered wavelengths (i.e., R: 1729.70nm, G: 1023.40nm, B: 752.43nm). Changes that do not affect these wavelengths are not visible in Fig.4-8.

4.4.2 Design of Experiments

The proposed S^2CVA approach was compared with other literature multiple change detection methods. First, the proposed S^2CVA and the standard C^2VA [47] representations were visually compared. Then change detection results were discussed through both a qualitative and a quantitative analysis based on the change reference map. The following techniques were considered for comparison: 1) standard automatic thresholding in the C^2VA feature space (C^2VA_T) [47]; 2) manual (interactive) change identification in the C^2VA feature space (C^2VA_M); 3) *k*-means clustering on the whole changed SCVs (*k*-means_SCVs); 4) proposed S^2CVA approach using manual (interactive) change identification at each level of the representation (S^2CVA_M). For the *k*-means_SCVs, the results are given as the average over 20 random initializations of the *k*-means algorithm. It worth noting that advantages were given to the *k*-means clustering by fixing the number of clusters (i.e., changes) as being known a priori (i.e., in the simulated cases the K is given as the number of the simulated changes, in the real data set K is given as the output of the proposed S^2CVA_M). EM algorithm was used in the framework of the Bayesian decision theory for estimating the thresholds in the C^2VA_T [47].

4.5 Experimental Results

4.5.1 Simulated Hyperspectral Remote Sensing Data Set

Fig.4-9 (a) shows the scattergram of all SCVs in the standard C^2VA feature space. In the binary CD step, threshold T_ρ was automatically estimated and resulted equal to 1.929 (see the red semicircle in Fig.4-9.a). Six change clusters can be observed from the C^2VA representation (despite eight were expected). The manual discrimination boundaries based on polygons were defined to separate them among each other and to extract the SCVs that correspond to each single cluster (see Fig.4-9.a, where the discrimination boundaries are defined in yellow polygons). From a comparison with the reference map, it was found that the correspondences among the six detected change clusters and the eight reference change classes are: ω_{C_6}, ω_{C_2}, $\omega_{C_1} \cup \omega_{C_7}$, $\omega_{C_3} \cup \omega_{C_4}$, ω_{C_5} and ω_{C_8}. Therefore, two clusters include different classes in the C^2VA representation. Moreover, some overlapped clusters (e.g., area between ω_{C_2} and ω_{C_6}, and the one between ω_{C_5} and ω_{C_8} in Fig.4-9.a) in the C^2VA representation cause detection errors, thus reducing the overall accuracy. If we separate the clusters with the discrimination boundaries reported in Fig.4-9 (a), the obtained CD map is shown in Fig.4-11 (b), where the detected six changes appear in different colors.

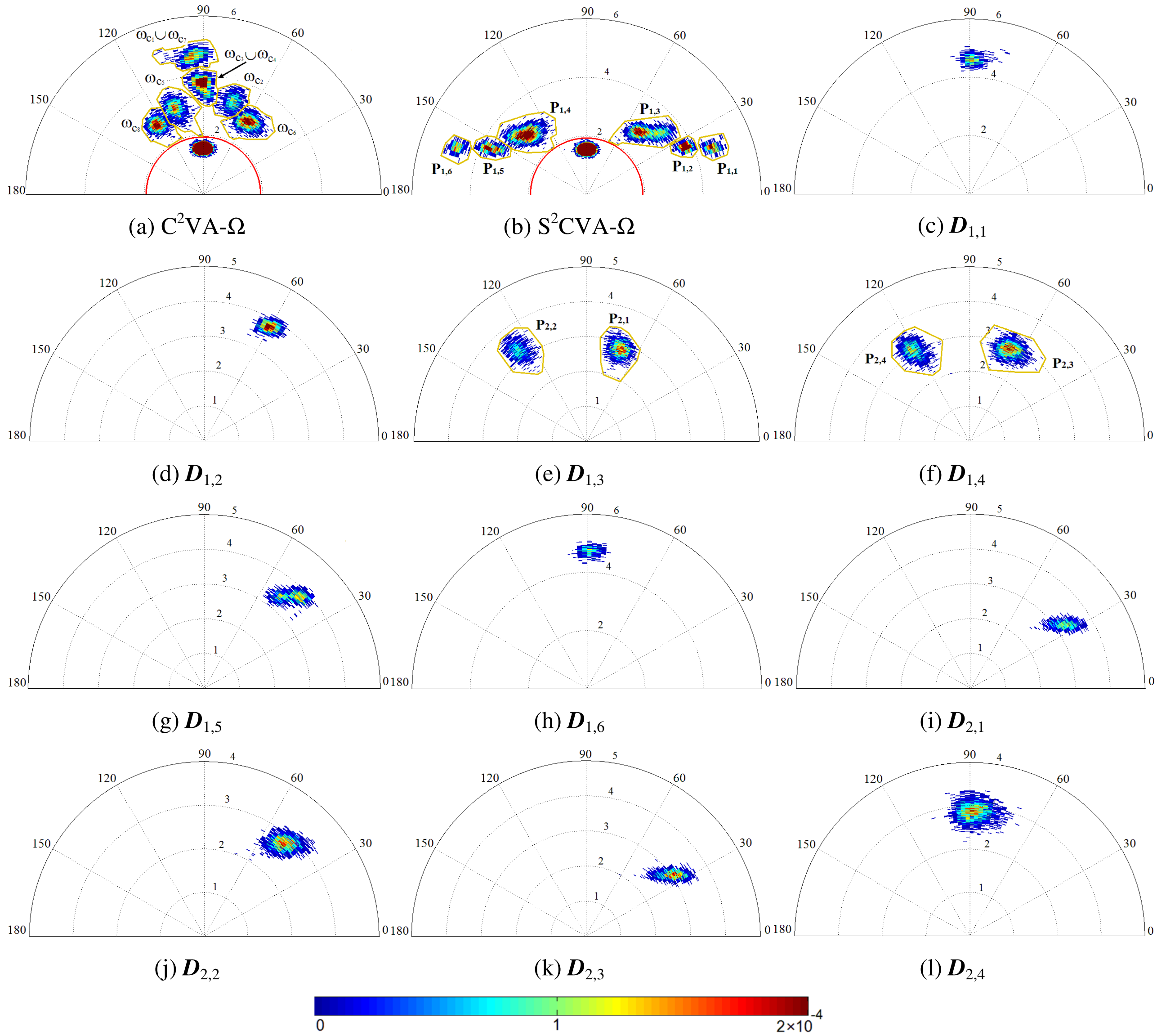


Fig.4-9 Change representation obtained by: (a) C^2VA_M; (b)-(l) proposed S^2CVA_M. The sequence of ASCVR scattergrams represents changes at different levels of the S^2CVA hierarchy. Binary CD decision threshold is defined as red semicircle, whereas discrimination boundaries as yellow polygons. The final detected change classes are those obtaining a single homogenous cluster in their corresponding representation scattergrams (simulated hyper-spectral remote sensing data set).

The same CD-HS task was addressed by using the proposed S^2CVA_M approach. We implemented it starting from the initial input Ω in L_0. The same threshold T_ρ estimated for C^2VA was used to separate Ω_c and ω_n (see the red semicircle in Fig.4-9.b). Six clusters inside of the general Ω_c class (P_0) were manually identified and separated into $P_{1,1}$, …, $P_{1,6}$ according to the defined discrimination boundaries in Fig.4-9 (b). Then the processing moved into the next level (i.e., L_1) and focused on each identified cluster $P_{1,1}$, …, $P_{1,6}$ to investigate the spectral homogeneity and further explore the possible presence of multiple changes in the re-projection in ASCVR scattergram. Reference vectors $\boldsymbol{R}_{1,j}$ (i.e., $\alpha_{1,j}$, j=1,…,6) were automatically derived for each considered specific cluster, thus generating a 2-D representation for each of them. The corresponding six represented scattergrams $\boldsymbol{D}_{1,1}$-$\boldsymbol{D}_{1,6}$ are illustrated in Fig.4-9 from (c) to (h). It is easy to see that four out of six clusters (i.e., $P_{1,1}$, $P_{1,2}$, $P_{1,5}$, $P_{1,6}$) resulted in single homogenous classes in their corresponding representation domains (see Fig.4-9.c-d, g-h). These homogenous clusters were associated to different detected change classes. The two discriminable clusters observed in the representation of $P_{1,3}$ and in $P_{1,4}$ (i.e., Fig.4-9.e and f, respectively) were further analyzed and separated in the next level L_2, where both of them appeared as a single homogenous cluster on the corresponding ASCVR scattergram $\boldsymbol{D}_{2,1}$-$\boldsymbol{D}_{2,4}$ (see Fig.4-9. i-l). Therefore, all eight change classes were successfully detected according to the proposed sequential analysis (four identified at level L_1 and four at level L_2).

The obtained hierarchal tree of the considered CD-HS problem is shown in Fig.4-10. The correspondences among the identified clusters and the reference change classes are as follows: L_0: $\Omega = \{\omega_n, \Omega_c\,(P_0)\}$; L_1: $P_{1,1} = \omega_{C_1}$, $P_{1,2} = \omega_{C_4}$, $P_{1,3} = \{\omega_{C_2}, \omega_{C_8}\}$, $P_{1,4} = \{\omega_{C_5}, \omega_{C_6}\}$, $P_{1,5} = \omega_{C_3}$ and $P_{1,6} = \omega_{C_7}$; L_2: $P_{2,1} = \omega_{C_2}$, $P_{2,2} = \omega_{C_8}$, $P_{2,3} = \omega_{C_5}$ and $P_{2,4} = \omega_{C_6}$.

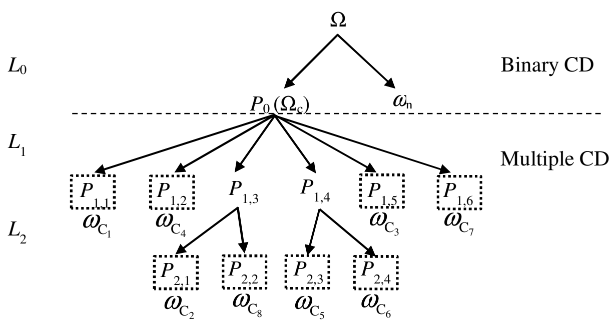


Fig.4-10 Three-level hierarchical tree obtained by the proposed S^{2}CVA_M method. Nodes highlighted in dotted rectangles are the final detected changes (simulated hyperspectral remote sensing data set).

From the qualitative comparison of the change representation between the proposed S^{2}CVA_M approach and the C^{2}VA_M method, we can observe that:

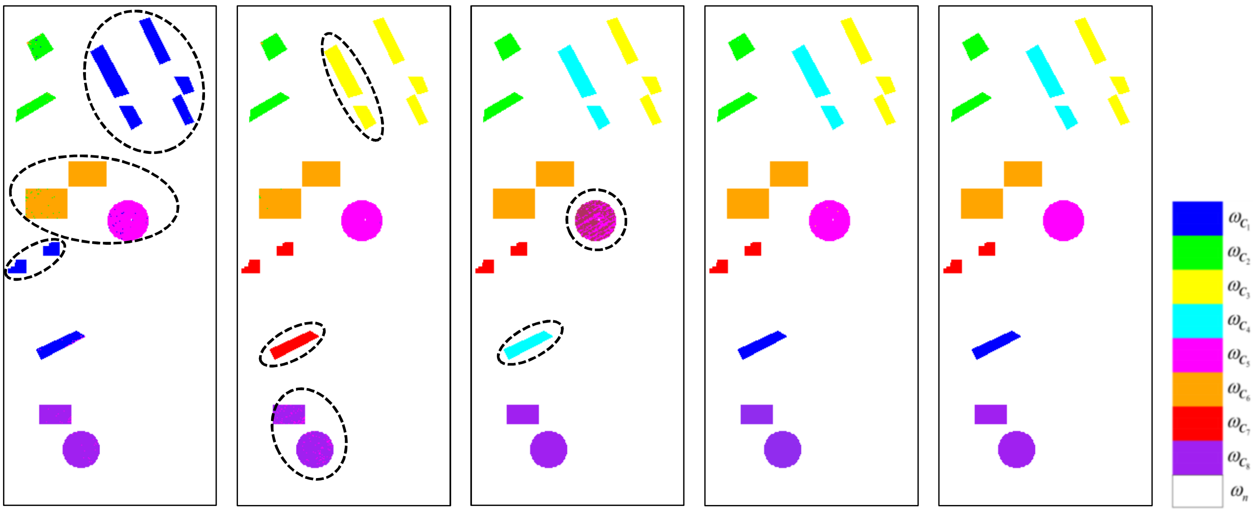
1) The proposed ASCVR resulted in an improved change representation. Higher class separability among the change clusters can be found at the initial level of the representation (Fig.4-9.b) than in C^{2}VA one (Fig.4-9.a). The inter-class distances among change clusters are larger, thus making it easier to discover the clusters associated with different changes and to define the discrimination boundaries. On the contrary in the C^{2}VA the change clusters are more compressed and overlapped (see Fig.4-9.a). This confirms the usefulness of choosing the maximum eigenvector as the adaptive reference vector.

2) The proposed S^{2}CVA_M method improves the change detectability. The hierarchical and adaptive scheme (i.e., use of adaptive reference vectors) for constructing spectral change variables allows one to discover and visualize more subtle spectral changes within the major change clusters detected at the first level L_1 of the hierarchy. A fixed reference vector like in the C^{2}VA does not permit to investigate the latent spectral variations at different spectral detail levels, thus in most of the cases only the major changes are identified.

3) The proposed S²CVA_M results in a better change modeling. As discussed in Section III, the complexity of CD-HS problem reduces the effectiveness of the single level processing of C²VA_M. As shown in Fig.4-9 (a), the C²VA representation does not distinguish some change classes. The proposed CD approach addresses the CD-HS problem by following a sequential fashion, thus it better considers the intrinsic hierarchical structure of changes in hyperspectral images.

TABLE 4-2 NUMBER OF DETECTED KINDS OF CHANGE, ACCURACY AND ERROR INDICES OBTAINED BY THE CONSIDERED METHODS (SIMULATED HYPERSPECTRAL REMOTE SENSING DATA SET).

CD Methods	**Number of detected kinds of change**	***OA* (%)**	***Kappa***	**Total errors (Number of pixels)**
C^2VA_T	5	-	-	-
C^2VA_M	6	-	-	-
k-means_SCVs	8	98.77	0.9256	1371
Proposed S^2CVA_M	8	99.99	0.9996	7



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

Fig.4-11 CD maps obtained by (a) C²VA_T; (b) C²VA_M; (c) *k*-means_SCVs; (d) proposed S²CVA_M. (e) change reference map. Different changes are in different colors, and the no-change class is in white (simulated hyperspectral remote sensing data set).

The qualitative and quantitative comparisons of the CD results obtained by the other considered reference methods are shown in Fig.4-11 and TABLE 4-2 (where accuracy indices include the Overall Accuracy (*OA*), Kappa Coefficient (*Kappa*) and the number of mislabeled samples computed according to the available reference map), respectively. The qualitative analysis was conducted only on the methods resulting in the correct number of changes (i.e., K=8). The highlighted dotted circles in Fig.4-11 point out the main omission/commission errors occurred in each of the considered method. From the analysis of the quantitative CD results, we can observe that the C²VA-based method did not detect all the changes due to the low class discriminability in the compressed change representation (see Fig.4-9.a). Only five and six changes were detected by using thresholding (i.e., C²VA_T) and interactive analysis (i.e., C²VA_M), respectively. Mixed changes and some false alarms are highlighted in Fig.4-11 (a) and (b). Although the *k*-means_SCVs was applied by providing as input the real number of changes (i.e., *K*=8), it resulted in a higher number of errors (i.e., 1371 pixels, mainly are commission errors, see Fig.4-11.c) than the proposed S^2CVA_M. The proposed S^2CVA_M approach achieved a very good CD result (see Fig.4-11.d, generated according to the sequence of scattergrams and the defined discrimination boundaries in Fig.4-9.b-l), resulting in the highest *OA* and *Kappa* values (i.e., 99.99% and 0.9996, respectively) with only 7 pixels of errors.

In addition, a detailed analysis of time taken from the proposed S^2CVA_M approach has been conducted. In the experiments we used Matlab R2013a on an Intel i5-2400 quad-core 3.10 GHz PC with 4 GB of RAM. Time consumption has been evaluated considering the time required for: i) obtaining the initial binary change-detection step (only for the root node); ii) running the 2-D ASCVR; iii) identifying changes (here the time for manual cluster separation is provided by an estimation based on an average of multiple users' trials). In this data set, the proposed S^2CVA required in total 166.96 seconds (less than 3 minutes) to complete the hierarchy, where the binary CD step took 31.59 seconds and the user interaction required around 120 seconds. For each of the nodes, the processing time (i.e., sum of the ASCVR and change identification) is in the range of [0.94, 2.05] seconds. Therefore, the computation cost is very low.

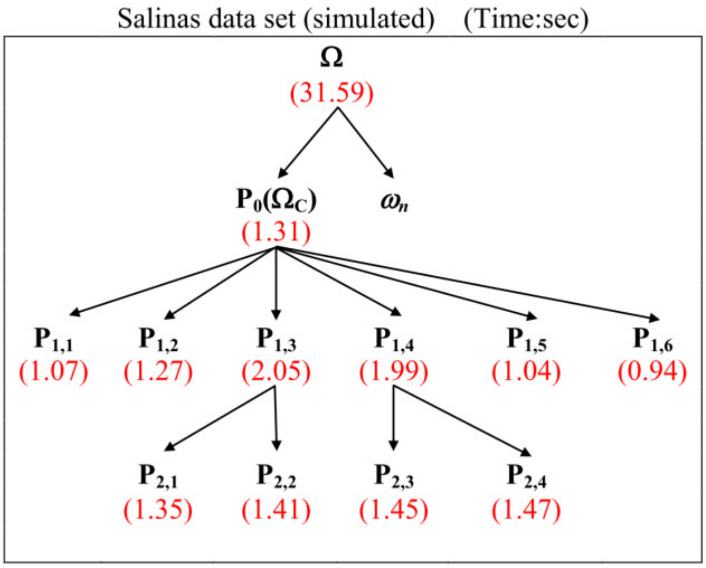


Fig.4-12 Summary of the computational time taken by the proposed technique (in second) on the Salinas simulated hyperspectral data set.

4.5.2 Simulated Hyperspectral Camera Data Set

The same reference CD methods considered in the previous case were also applied to this data set. T_ρ for binary CD step was equal to 0.278. The C^2VA and S^2CVA representations are shown in Fig.4-13 (a) and (b)-(o), respectively, where the interactive change identification was done by the defined boundaries (see Fig.4-13). In this case the reference change map is also available due to the simulation procedure, so a fully quantitative evaluation was done by comparing the results obtained by the considered CD methods as in the previous experiment. The numeric results are shown in TABLE 4-3.

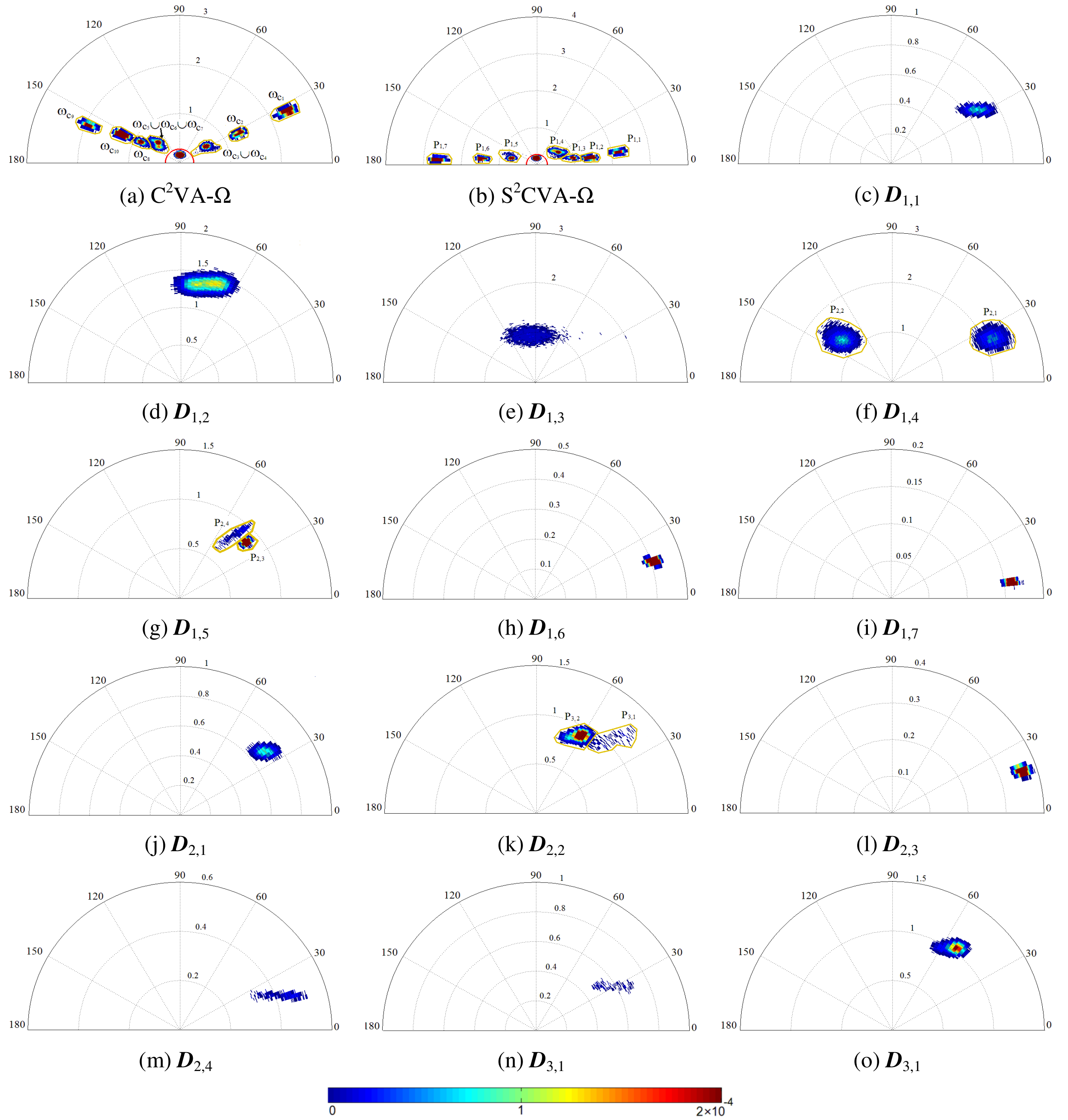


Fig.4-13 Change representation obtained by: (a) the C^{2}VA_M; (b)-(o) the proposed S^{2}CVA_M. The sequence of ASCVR scattergrams represents changes at different levels of the S^{2}CVA hierarchy. Binary CD decision threshold is defined as red semicircle, whereas discrimination boundaries as yellow polygons. The final detected change classes are those obtaining a single homogenous cluster in their corresponding representation scattergrams (simulated hyperspectral camera data set).

TABLE 4-3 NUMBER OF DETECTED KINDS OF CHANGE, DETECTION ACCURACY AND ERROR INDICES OBTAINED BY THE CONSIDERED METHODS (SIMULATED HYPERSPECTRAL CAMERA DATA SET).

CD Methods	**Number of detected kinds of change**	***OA* (%)**	***Kappa***	**Total errors (number of pixels)**
C^2VA_T	4	-	-	-
C^2VA_M	7	-	-	-
k-means_SCVs	10	98.09	0.8949	12791
S^2CVA_M	10	99.94	0.9964	801

From the results in TABLE 4-3, we can observe that:

1) The C^2VA_M separated and detected more changes (i.e., *K*=7) than the C^2VA_T (i.e., *K*=4). However, both of them recognized less changes than the correct one (i.e., *K*=10) due to the poor change representation in C^2VA using the fixed unit reference vector.

2) The proposed S^2CVA_M detected correctly all the ten changes, and achieved the highest *OA* (i.e., 99.94%) and *Kappa* (i.e., 0.9964). In this case the top-down procedure resulted in the four-level hierarchical tree as shown in Fig.4-14. Moreover, a fast and simple change identification was done interactively by using the proposed S^2CVA_M approach.

3) The proposed S^2CVA_M modeled better the hierarchical nature of the changes in hyperspectral images thus reduced the detection errors. Note that also in this case the *k*-means_SCVs method was not able to correctly detect all changes even if it received as input the correct number of changes (i.e., *K*=10).

The same computation-cost evaluation was conducted as for the previous data set. The proposed S^2CVA_M took 437.13 seconds (less than 8 minutes), where the binary step and manual interaction required 317.13 and 120 seconds, respectively. For each node, the processing cost is in the range of [4.72, 11.62] seconds. Thus, despite this data set has a large size (i.e., 1392×1040×31), the proposed S^2CVA_M still resulted in a low computation cost.

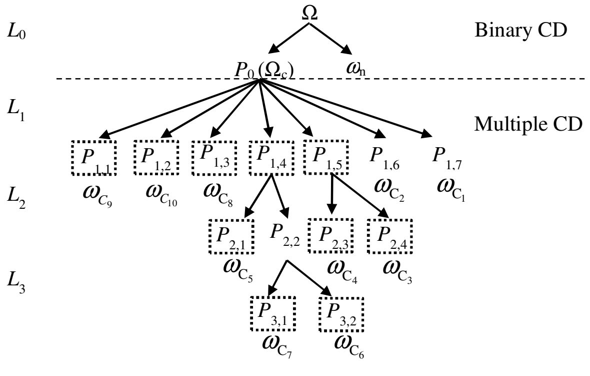


Fig.4-14 Four-level hierarchical tree obtained by the proposed S^2CVA_M method. The nodes highlighted in dotted rectangles are the final detected changes (simulated hyperspectral camera data set).

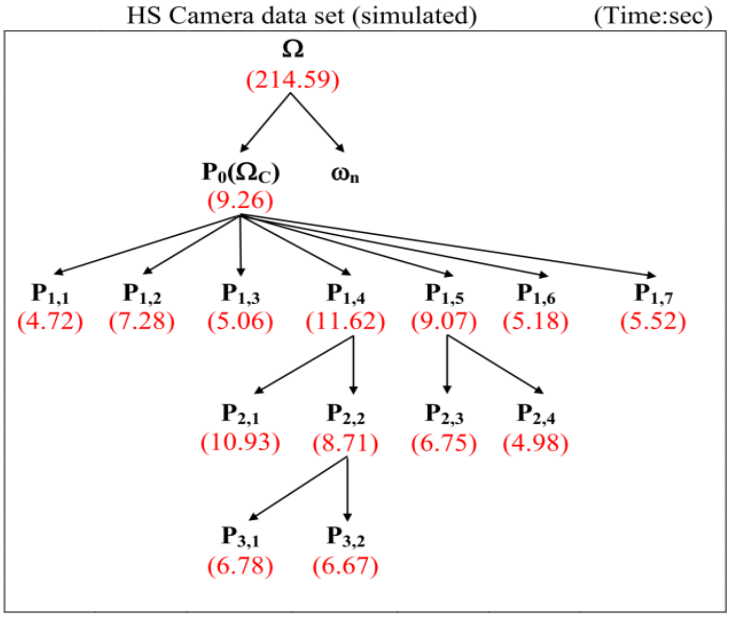


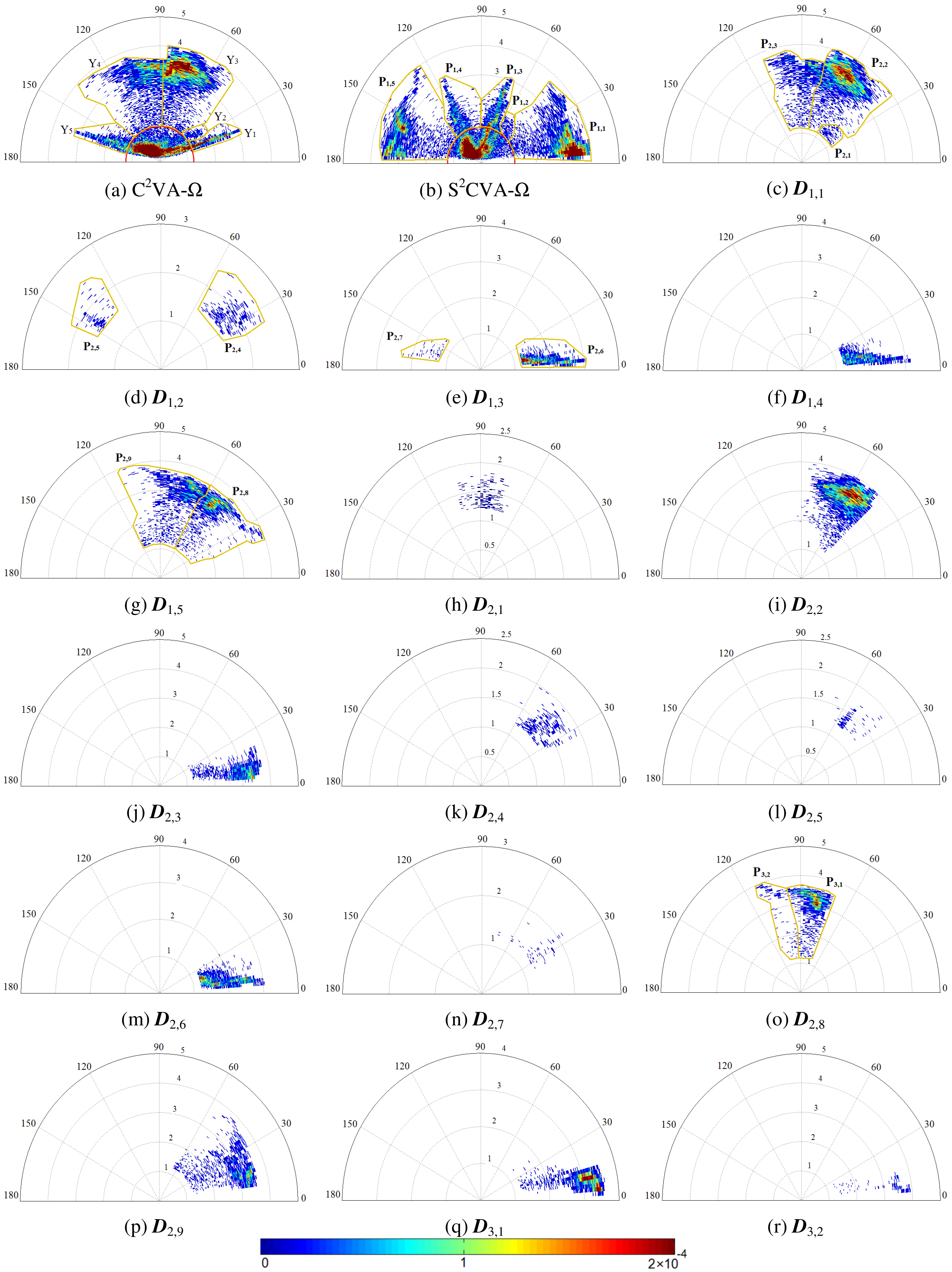
Fig.4-15 Summary of the computational time taken by the proposed technique (in second) on the simulated hyper-spectral camera data set.

4.5.3 Real Hyperion Remote Sensing Satellite Data Set

Both the standard C^2VA and the proposed S^2CVA methods were applied to the considered images. The representation obtained by the C^2VA is shown in Fig.4-16 (a). The binary CD step was conducted on the magnitude ρ computed from all 159 spectral channels. The threshold T_ρ for separating the Ω_c and ω_n was estimated automatically [31] and was equal to 1.219 (see Fig.4-16.a, where T_ρ is represented as a red sem-icircle). Change identification was conducted, where the boundaries were interactively defined as yellow polygons. Five different change classes (see Fig.4-16.a) were detected. For the S^2CVA, the whole change structure and the sequence of ASCVR scattergrams obtained are illustrated in Fig.4-16 from (b) to (r), where each figure corresponds to a specific change cluster that is represented in a given level of the S^2CVA hierarchy. In the initial level of the sequential analysis (i.e., L_0), the binary CD step was per-formed as in the C^2VA (see Fig.4-16.b). Then the SCVs of the Ω_c class (P_0) were analyzed to discover other change classes and discrimination boundaries were defined for each of the scattering cluster in the representation domain (see Fig.4-16.b, $\boldsymbol{D}_{1,1}$-$\boldsymbol{D}_{1,5}$). The S^2CVA process iterated until convergence was reached in each level of the hierarchy. Eleven kinds of change were detected by using the proposed S^2CVA method. The hierarchical tree is described in Fig.4-17. The obtained tree has four levels with nineteen nodes. Eleven of them correspond to the detected change classes and one to the no-change class.

By analyzing the results shown in Fig.4-16 and Fig.4-17, we can observe that:

1) The C^2VA representation (i.e., Fig.4-16.a) results in changes that are overlapped to each other and their discrimination is almost impossible. A smaller number of change classes is identified by C^2VA_M (i.e., K=5, see Fig.4-16.a, cluster Y_1-Y_5) than by the proposed S^2CVA_M approach (i.e., K=11).



(a) C^2VA-Ω (b) S^2CVA-Ω (c) $\boldsymbol{D}_{1,1}$

(d) $\boldsymbol{D}_{1,2}$ (e) $\boldsymbol{D}_{1,3}$ (f) $\boldsymbol{D}_{1,4}$

(g) $\boldsymbol{D}_{1,5}$ (h) $\boldsymbol{D}_{2,1}$ (i) $\boldsymbol{D}_{2,2}$

(j) $\boldsymbol{D}_{2,3}$ (k) $\boldsymbol{D}_{2,4}$ (l) $\boldsymbol{D}_{2,5}$

(m) $\boldsymbol{D}_{2,6}$ (n) $\boldsymbol{D}_{2,7}$ (o) $\boldsymbol{D}_{2,8}$

(p) $\boldsymbol{D}_{2,9}$ (q) $\boldsymbol{D}_{3,1}$ (r) $\boldsymbol{D}_{3,2}$

Fig.4-16 Change representations obtained by: (a) C^2VA_M; (b)-(r) the proposed S^2CVA_M approach. A sequence of ASCVR scattergrams represents changes at different levels of the S^2CVA hierarchy. Binary CD threshold (i.e., T_ρ) is defined as red semicircle, whereas discrimination boundaries are defined as yellow polygons. The final detected change classes are those obtaining single homogenous cluster in their corresponding scattergrams (real hyperspectral Hyperion remote sensing data set).

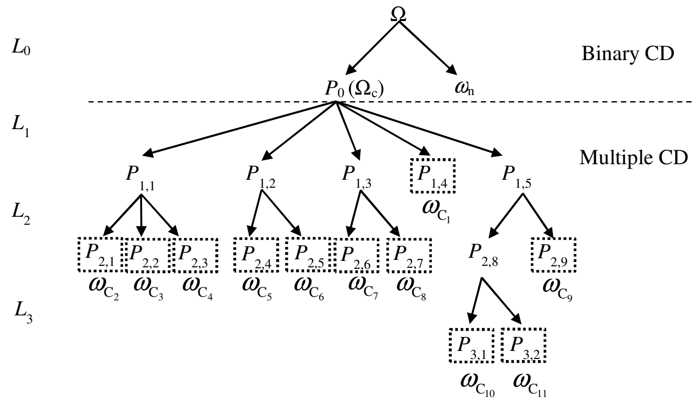


Fig.4-17 Four-level hierarchical tree obtained by proposed S^2CVA_M method. Eleven change classes were identified and highlighted in dotted rectangles (real Hyperion remote sensing data set).

2) The proposed S^2CVA_M method successfully addressed the considered CD-HS problem by decomposing it into several sub-problems. The multiple-change information was modeled and represented through seventeen ASCVR scattergrams. The sequence is automatically defined by investigating the inner-cluster spectral homogeneity at different levels. Note that changes merged in low level representations become visible and separable in higher levels, thus allowing us to detect different kinds of changes in a hierarchical way, whose structure is described by the hierarchy tree in Fig.4-17.

3) The proposed unsupervised ASCVR approach properly discovers and represents changes by taking advantages of the specific SCVs for defining the reference vectors thus for computing the change variables. For each considered change cluster at a given level, the homogeneity is evaluated according to the represented 2-Dimensional scattergram, where a pure change results in a single homogenous scattering cluster. For example, $\boldsymbol{D}_{1,3}$ represents two discriminable clusters in cluster $P_{1,3}$, whereas $\boldsymbol{D}_{1,4}$ represents only one cluster in $P_{1,4}$, which indicates a higher homogeneity of the latter.

It is worth noting that the proposed method allows users to control the detection level, which is a very flexible property in real applications. If the aim is to identify more detailed subtle-change information, the process can go deeper down the hierarchy (i.e., detection of subtle changes that have slightly different realizations on their SCVs from the spectral behaviors of an associated change in the previous level [86]). In the proposed S^2CVA hierarchy, a change class is detected when it is associated with one homogenous cluster in its ASCVR representation (e.g., Fig.4-16.f, h-n, p-r). Sometime small changes (e.g., less than 10 pixels) might be discovered in the representation. For example, in Fig.4-16 (k), (n), (p) there are some SCVs that are isolated from the main (lager) change clusters. These kinds of clusters can be relevant to some applications whereas they can be irrelevant to others. Thus the user can decide to stop the analysis at higher levels in the hierarchy where those changes are not separated yet. In our experiments, we followed this strategy and did not separate them.

The final CD maps generated by the proposed S^2CVA_M and the considered reference methods are shown in Fig.4-19 (a)-(d). Different change classes appear in different colors and the no-change class is in white. The number of the detected changes in each method is listed in TABLE 4-4. Note that also in this case advantages were given to *k*-means_SCVs method providing as input the number of changes (i.e., K=11) obtained by the proposed S^2CVA_M. The performances of the considered methods were analyzed in a qualitative way by a detailed visual comparison.

TABLE 4-4 NUMBER OF DETECTED KINDS OF CHANGE IN THE CONSIDERED METHODS (REAL HYPERION REMOTE SENSING DATA SET).

CD Methods	**Number of detected kinds of change**
C^2VA_T	5
C^2VA_M	5
k-means_SCVs	11
S^2CVA_M	11

As we can see from Fig.4-19 and TABLE 4-4, the two C^2VA-based methods (i.e., C^2VA_T and C^2VA_M) resulted in a small number of changes (i.e., K=5), which correspond to the major change classes. These changes show significant differences in spectral signatures in the SCV domain. However, many subtle changes are not visible and thus not detectable due to the compressed representation in C^2VA by using a fixed unit reference vector. The C^2VA_M considered better the distribution of changes (Fig.4-16.a) thus resulting in a more reliable CD output than the C^2VA_T (see Fig.4-19.b than Fig.4-19.a). By using the iterative analysis of the proposed S^2CVA method, more subtle changes became visible and detectable (see Fig.4-19.f) according to a systematic top-down procedure. These subtle change classes were not detectable by the C^2VA-based methods (see Fig.4-19.a-b).

The two C^2VA-based approaches detected fewer changes than the S^2CVA_M. Some latent changes are still mixed in some detected clusters, which cannot be discriminated and separated by using C^2VA (see Fig.4-19.a and b). The S^2CVA_M provided more convincing results than the *k*-means_SCVs. This can be observed in Fig.4-19.d, where more homogenous change classes are present in the change-detection map. On the contrary, the *k*-means_SCVs resulted in more fragments in the detected changes (see Fig.4-19.c). It is worth noting that the complex structure of the problem makes the use of clustering (which solves an ill-posed problem) less reliable than that of a user-manual decision. In the S^2CVA_M users can control the decomposition into different levels, and the whole change representation and identification process.

From the computation-cost point of view, the proposed S^2CVA_M resulted in a very fast and efficient implementation. In this data set, S^2CVA_M took in total 177.99 seconds (less than 3 minutes) to evaluate the four-level hierarchy for the change discovery, representation and detection. In greater details, the binary CD step took 24.80 seconds, and the user interaction required around 120 seconds.

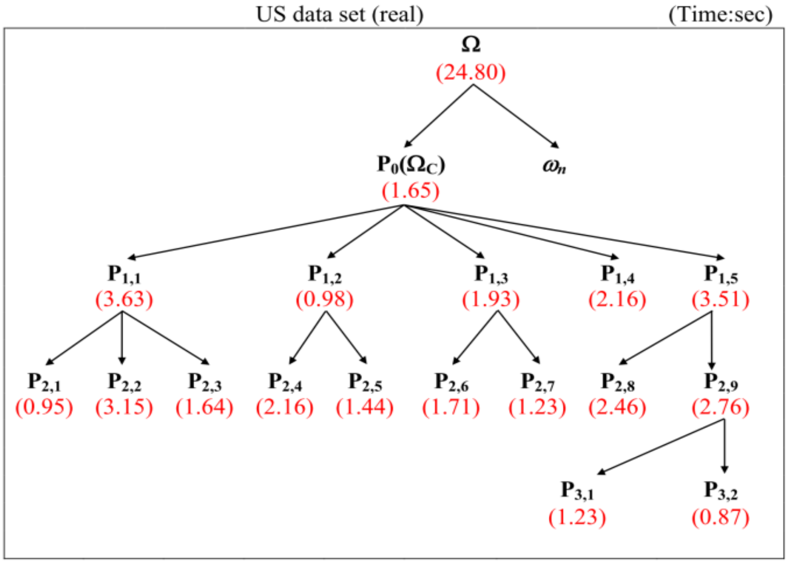


Fig.4-18 Summary of the computational time taken by the proposed technique (in second) on the real Hyperion hyperspectral data set.

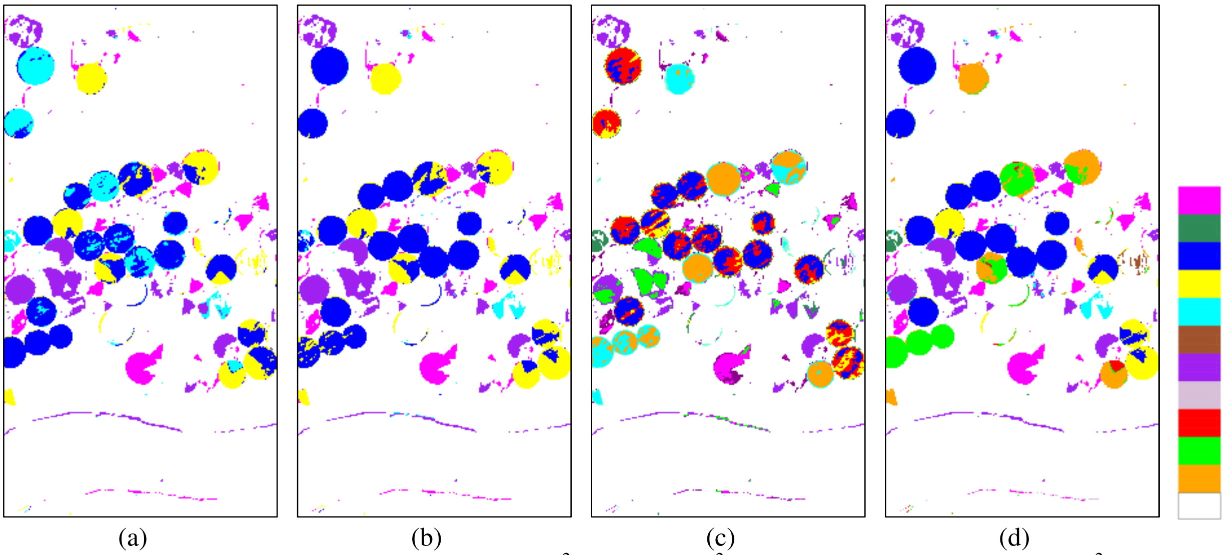


Fig.4-19 Change-detection maps obtained by (a) C^2VA_T; (b) C^2VA_M; (c) *k*-means_SCVs; (d) S^2CVA_M (according to the discrimination boundaries defined in Fig.4-16. (b)-(e), (g) and (o)). Different changes are in different colors, and the no-change class is in white color (real hyperspectral Hyperion remote sensing data set).

4.6 Discussion and Conclusions

In this chapter, a novel sequential spectral change vector analysis (S^2CVA) approach has been proposed in order to address the challenging multiple-change detection problem in multitemporal hyperspectral images. Developed on the basis of the C^2VA state-of-the-art method, the proposed approach aims at discovering, representing and detecting multiple changes according to a sequential process that takes into account different levels of spectral change significance. The main novelties of the proposed S^2CVA are: 1) It iteratively analyzes the heterogeneous change information by following a top-down structure and a sequential analysis. Thus changes can be represented, discovered and detected at different levels of the hierarchy. 2) At each level it adaptively exploits the proposed ASCVR to represent changes by using a reference vector automatically and adaptively defined. Experimental results obtained on both simulated and real multitemporal hyperspectral images confirmed the effectiveness of the proposed approach.

Based on the theoretical analysis and the empirical experimental results we can also conclude that:

1) The proposed S^2CVA method extends the use of C^2VA to hyperspectral images where a large number of major and subtle changes may be present in the high dimensional data. Changes are discovered and separated according to their intrinsic spectral behaviors in SCVs, which are represented by a hierarchical tree. The computational complexity of the proposed S^2CVA method is very low (in all our experiments few minutes were required for the entire processing on a standard PC). It is worth noting that, despite the total processing time depends on the hierarchical tree size, any additional node requires in average few seconds. Thus the iterative nature on the process does not represent a critical limitation in real applications. Note that the proposed method can also be used for addressing the CD-MS problem. In this case we expect that the hierarchy results in fewer levels.

2) Unlike the standard C^2VA, the proposed ASCVR method adaptively and automatically changes the reference vectors according to the SCVs of the specific changes that are analyzed. Therefore, although the compression from the B-Dimensional to the 2-Dimensional feature space introduces an unavoidable loss of information, the sequential analysis gradually recovers in the hierarchy the information loss at first levels, resulting in complete change representations and in satisfactory change-detection maps.

3) The proposed change identification approach allows us to directly extract the change information in the ASCVR domain of interest of the user, thus providing an easy but efficient way to address the change discovery and separation problem in complex problems with hyperspectral images.

An apparent limitation of the proposed method is that it results in a semi-automatic implementation (changes are detected via interaction with the user), which does not allow a completely automatic detection of changes. However, we would like to point out that the main goal of the proposed approach is to have an effective top-down procedure that supports the user in discovering and analyzing changes through an interactive process. This is very important for addressing CD problems with hyperspectral images.

Future developments of this work will be focused on: 1) study of other change representation variables aimed to further enhance the change representation; and 2) joint use of the spatial-spectral multiresolution information to reduce the misregistration effect.

Chapter 5

A Novel Hierarchical Clustering Method for Change Detection in Hyperspectral Images

Multitemporal hyperspectral images provide very detailed spectral information that directly relates to land surface composition. This results in the potential detection of more spectral changes than those visible in the traditional multispectral images. However, the process of automatically extracting changes from hyperspectral images is very complex. This chapter addresses the multiple-change detection problem in multitemporal hyperspectral remote sensing images by analyzing the complexity of this task. A novel partially-unsupervised hierarchical change-detection approach is presented, which aims to identify the possible changes occurred between a pair of hyperspectral images based on a designed hierarchical spectral change clustering. Changes having discriminable spectral behaviors in hyperspectral images are identified hierarchically by following a top-down structure. A manual initialization is used to trigger the clustering, whereas the clustering itself is totally unsupervised. Experimental results obtained on simulated and real bi-temporal images confirm the validity of the proposed hierarchical change detection approach.

5.1 Introduction

In this chapter we focus the attention on effective clustering methods that exploit the difference image $\boldsymbol{X}_D$. The difference image (computed by subtracting pixel by pixel in spectral channels) carries multiple change information. Thus the behavior of SCV signatures in the hyperspectral difference domain $\boldsymbol{X}_D$ results in a fine modeling of different kinds of changes, which is not possible with multispectral images. To better understand this concept, let us consider a vegetated field affected by land-cover changes. On the one hand, multispectral images can highlight strong changes, which are class transitions that significantly affect the spectral signature (e.g., vegetation to land covers like water, built-up areas, soil). However within such strong changes, other changes may be observed that correspond to slightly different realizations of the strong change itself. In a given vegetation change class there might be more change contributions due to different factors (e.g., difference on the vegetation growth status, density, water content). These kinds of changes show small spectral differences with respect to those of the strong change they are associated with. Such differences are usually localized in specific parts of the spectrum, which are usually difficult to be recognized from the rough spectral representation in typical of multispectral images. On the

This chapter is published on *IEEE Transaction on Geoscience and Remote Sensing*, Vol. 53, no. 1, 2015, pp. 244-260. Title: "Hierarchical Unsupervised Change Detection in Multitemporal Hyperspectral Images". Authors: S. Liu, L.Bruzzone, F. Bovolo, P.Du.

other hand, these subtle changes become detectable in hyperspectral images due to the detailed representation of the spectral signatures. Moreover, if calibrated data are available, it is possible to obtain the explicit semantic meaning of the class transition ("from-to") for a change by matching the spectral signature of each single date with the standard reference spectra in spectral libraries. However, in the reality, reference samples are often not available. Therefore, the design of effective unsupervised CD methods that are independent from ground truth data availability is highly attractive in real applications.

In this chapter, a hierarchical partially-unsupervised CD approach is present that is suitable to identify different kinds of changes between two hyperspectral images [86]. The developed CD method: 1) addresses the problem of multiple-change detection; 2) makes adequate use of the detailed spectral information in hyperspectral data; and 3) is partially-unsupervised.

The outline of this chapter is as follows. The proposed CD method based on the hierarchical clustering is described in detail in Section 5.2. The used remote sensing data sets and the experimental setup are introduced in Section 5.3. The experimental results are shown and discussed in Section 5.4. Finally, Section 5.5 draws the conclusion.

5.2 Proposed Approach

Based on the discussion, definitions and assumptions presented in Chapter 3, Section 3.1, the aim of the proposed approach is to detect the change endmembers in $\Omega_e=\{e_1, e_2,\ldots,e_K\}$ that each of them associated to a given change class in $\Omega_c=\{\omega_{c_1},\omega_{c_2},\ldots,\omega_{c_K}\}$. We propose a novel hierarchical CD method for detecting changes in hyperspectral images and separating them into different change endmembers. The proposed method mainly consists of three steps: a) pseudo-binary change detection to initialize the process and extract general changes; b) change endmember detection based on hierarchical spectral change analysis; and c) generation of the CD map by merging endmember clusters. The block scheme of the proposed approach is illustrated in Fig.5-1.

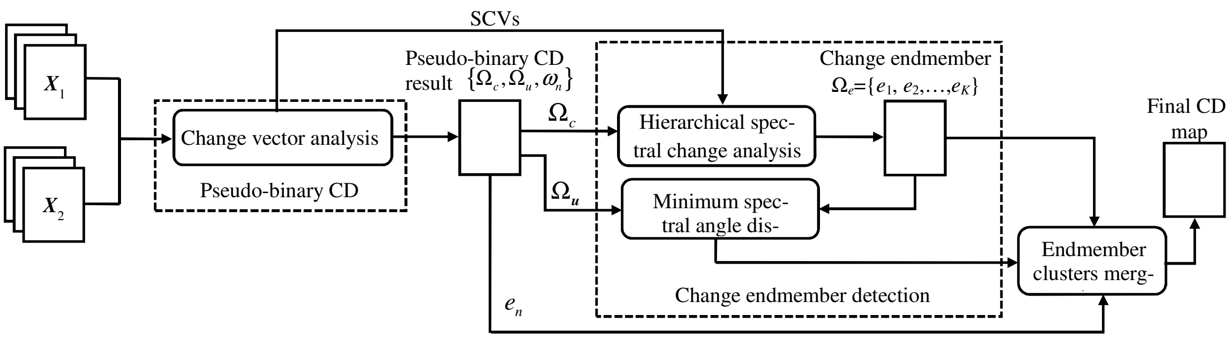


Fig.5-1 Block scheme of the proposed change-detection approach to multitemporal hyperspectral images.

5.2.1 Pseudo-Binary Change Detection

This step is based on the analysis of the magnitude of SCVs according to traditional binary CD techniques. However it is referred as pseudo-binary because the output has three classes. After separating the change (Ω_c) and no-change (ω_n) classes (thus no-change endmember e_n is straightforward), an uncertainty buffer class (Ω_u) is defined. The class of changes (Ω_c) is used to initialize the root node of a tree structure for change representation.

From $\boldsymbol{X}_D$ the magnitude and the direction of SCVs can be extracted. In the first step of the proposed method we are only interested in distinguishing Ω_c from ω_n. Thus only the magnitude ρ is considered as defined in (6). Thus the whole B-Dimensional change information is compressed into a 1-Dimensional feature. The rationale behind this choice is: 1) to simplify and avoid any feature selection procedure; 2) to exploit the contribution of all portions of the spectrum. If noisy bands are detected in the pre-processing (e.g., due to atmosphere absorption) they can be neglected.

Changed and unchanged pixels are separated into two groups according to a threshold value T_ρ computed on the magnitude variable. The Bayesian decision theory is applied to find this threshold [31]. The Expectation Maximization (EM) algorithm is used for estimating the class statistical parameters (i.e., the class prior probabilities, the mean values and variances) in an unsupervised way [31], [103]. Note that change and no-change classes are assumed to be Gaussian distributed, and multiple changes are approximated as one single change class (Ω_c) in the magnitude domain to focus only on the general change information. This approach has been widely used in binary CD with multispectral images and demonstrated to be a good approximation in hyperspectral images [6], [33], [47]. The approximation is acceptable as this is only a preliminary step.

In order to reduce the effect of possible thresholding errors and obtain conservative results that do not propagate significant errors in the next steps, a margin δ is set on the threshold computed on the histogram $h(\rho)$ of the magnitude ρ (see Fig.5-2) and three classes are defined. The three classes are: 1) *class of uncertain pixels* (Ω_u), on which it is not possible to take a reliable decision at this level of the processing. These pixels will be analyzed and reclassified according to the generated endmembers; 2) *class of changed pixels* (Ω_c), which includes pixels having a high probability to be changed, but without any information on their kind. The problem of the multiple changes identification will be addressed in the next step by the proposed hierarchical spectral change analysis method; 3) *class of no-changed pixels* (ω_n), which only contains pixels having a high probability to be unchanged. These pixels are treated as a pure no-change class endmember due to their low magnitude. Thus for a given SCV $\boldsymbol{x}(i, j)$ in $\boldsymbol{X}_D$ ($1 \le i \le I$, $1 \le j \le J$), a label is assigned according to the following rule:

$$\boldsymbol{x}(i,j) \in \begin{cases} \Omega_c, & if \ \ \rho(i,j) \ge T_\rho \\ \Omega_u, & if \ \ T_\rho - \delta \le \rho(i,j) < T_\rho \\ \omega_n, & if \ \ \rho(i,j) < T_\rho - \delta \end{cases} \tag{18}$$

where $\rho(i,j)$ is the SCV magnitude of the considered $\boldsymbol{x}(i, j)$. Fig.5-2 illustrates the flowchart of the pseudo-binary CD step.

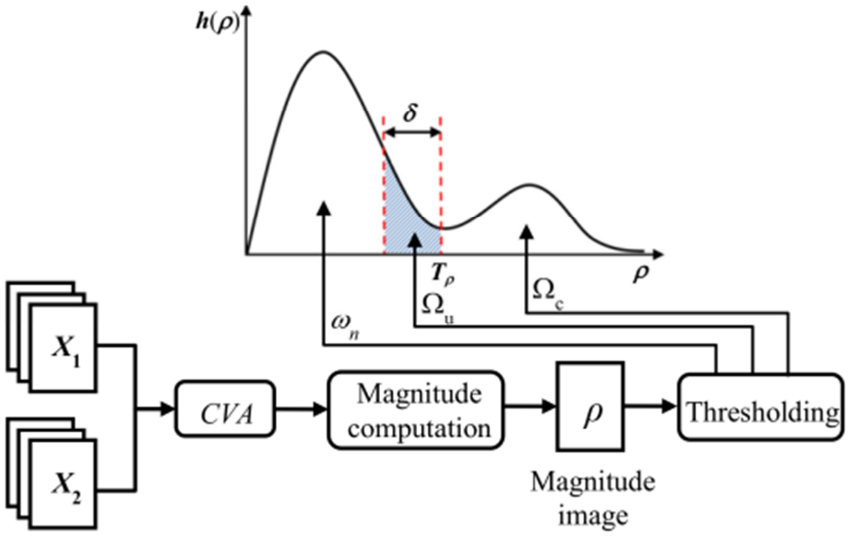


Fig.5-2 Block scheme of the pseudo-binary change-detection step used for initializing the tree structure.

5.2.2 Hierarchical Spectral Change Vector Analysis (HSCVA)

Let us focus on the classes of changed (i.e., Ω_c) and uncertain (i.e., Ω_u) pixels obtained in the previous step for identifying the change endmembers. The problem can be addressed by using clustering methods to automatically find the different change classes. However, the problem of multiple-class separation in hyperspectral images is much more difficult than in multispectral images. This is due to the following issues: 1) the high spectral resolution makes the spectrum more sensitive to changes, thus a high number of changes might be detected; and 2) subtle changes within major changes are always difficult to be identified directly from Ω_c. These problems decrease the detectability of all the hierarchy of changes directly from the data in one shot, and limit the effectiveness of clustering methods.

To overcome the mentioned problems, we propose a solution based on the idea of decomposing the original complex problem into sub-problems by a *Hierarchical Spectral Change Vector Analysis* (HSCVA) (see Fig.5-3 for a qualitative example of hierarchy). The hierarchical structure is modeled by a tree of changes defined to drive the analysis. Let L_d be a generic level in the tree structure with d = 0, 1,…, D-1. The depth of the tree is D (e.g., D=4 in Fig.5-3). The main idea is to start from the root node in the top level (i.e., L_0 that represents the general change class Ω_c identified in the pseudo-binary CD step) and gradually separate different kinds of change into child nodes by selectively exploiting the spectral information. At the first level (i.e., L_1) of the tree the priority is given to identify the major changes that according to the definition of Chapter 3, Section 1 have significant spectral difference from each other. Within each child node, subtle changes (if any) are detected and separated. This process is iterated until all change endmembers (i.e., leaves of the tree) are found.

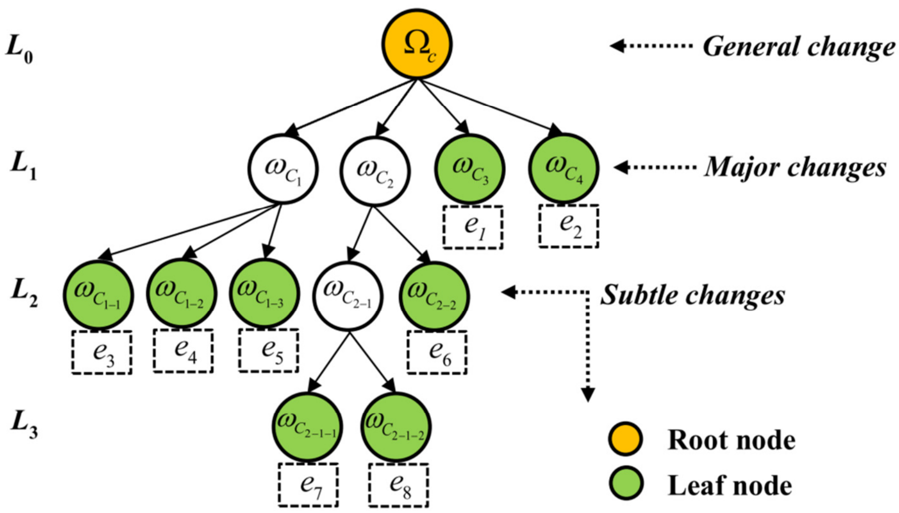


Fig.5-3 Example of the proposed hierarchical tree for the detection of change endmembers with tree depth D=4 and eight identified leaves.

Let us consider the root node that contains all the changed pixels without any distinction about their kind. To model the spectral homogeneity of Ω_c, a similarity measure based on the Spectral Angle Distance (SAD) [104] is used. The SAD ϑ is computed between each $\boldsymbol{x}(i,j)$ in Ω_c, and a reference spectral signature $\boldsymbol{S}_{\Omega_c}$ calculated as the average of all the $\boldsymbol{x}(i,j)$ in Ω_c, i.e.,

$$\vartheta(\boldsymbol{x}(i,j),\boldsymbol{S}_{\Omega_c})=\arccos\left(\frac{\sum_{b=1}^{B}\boldsymbol{x}^b(i,j)\boldsymbol{S}^b_{\Omega_c}}{\sqrt{\sum_{b=1}^{B}\left(\boldsymbol{x}^b(i,j)\right)^2}\sqrt{\sum_{b=1}^{B}(\boldsymbol{S}^b_{\Omega_c})^2}}\right) \tag{19}$$

where $\boldsymbol{x}^b(i,j)$ and $\boldsymbol{S}^b_{\Omega_c}$ are the b-th component in $\boldsymbol{x}(i,j)$ and $\boldsymbol{S}_{\Omega_c}$, respectively. For each $\boldsymbol{x}(i,j)$, the smaller $\vartheta(\boldsymbol{x}(i,j),\boldsymbol{S}_{\Omega_c})$, the higher the similarity with the reference spectrum and vice versa. For a pure change endmember we expect that all SCVs have very similar spectral behaviors, thus resulting in a small standard deviation of the similarity measure. Thus to verify the homogeneity of Ω_c we compare the standard deviation value $\sigma_{\vartheta_{\Omega_c}}$ of $\vartheta(\boldsymbol{x}(i,j),\boldsymbol{S}_{\Omega_c})$ with a predefined threshold value T_σ. If $\sigma_{\vartheta_{\Omega_c}}$ is smaller than T_σ, the change class is considered as being homogeneous and a change endmember is detected. Accordingly, the process is in convergence and the tree only has a single node. Otherwise the change class is considered as being inhomogeneous and likely to contain more than one kind of change. Therefore the hierarchical decomposition starts.

To distinguish major changes in Ω_c the Principal Component Analysis (PCA) and clustering algorithm are used. However, any other transformation technique can be considered. Note that PCA is applied only to the SCVs belonging to Ω_c. In this way we optimize the representation of the changes. Then the clustering algorithm is applied to the subset of transformed Principal Components (PCs) that includes more than 95% of change information to reject the noise and redundant information. This choice also reduces the computational complexity. Let $\boldsymbol{Q}$ be the image with selected M ($M < B$) PCs and let $\boldsymbol{Q}(i,j)$ be the vector characterizing spatial position (i,j) in $\boldsymbol{Q}$, $\boldsymbol{Q}(i,j)\in\boldsymbol{Q}$. An effective clustering technique should be used to correctly identify the major change classes inside Ω_c. The following issues need to be addressed: 1) identification of the number of major changes; 2) definition of a strategy for modeling and clustering the change information.

In order to address the above two issues, the adaptive x-means algorithm is used to automatically find an optimal number of major changes and generate reliable clustering results in an unsupervised framework [105], [106]. Differently from the popular k-means method, x-means adaptively searches on a range of k values and finds the best clustering model according to the Bayesian Information Criterion (BIC) [105]. The BIC identifies an adequate tradeoff between simplicity of the model (number of parameters) and quality of fit. It analyzes the maximum likelihood-based models of a given data distribution. We adopt the algorithm proposed in [106], which is an expansion of the original x-means, and modified it in order to satisfy our requirements. A given range G = [k_0, k_0+t] is first defined to initialize the x-means. This is the only input parameter to the algorithm. k_0 denotes the lower bound for the number of major changes k, and t is a constant value to control the upper bound. Then M-Dimensional PCs of Ω_c are given as input to the x-means clustering and the method is initialized by applying conventional k-means with k = k_0. We assume that all kinds of change approximately follow the Gaussian distribution.

For a given class ω_{c_k}, let $\boldsymbol{Q}_k$ be the pixel data of PCs, whose probability density function $f(\cdot)$ can be written as:

$$f\left(\boldsymbol{\Theta}_k,\boldsymbol{Q}_k\right)=(2\pi)^{-M/2}\left|\boldsymbol{\Gamma}_k\right|^{-1/2}\times\exp\left[-\frac{1}{2}(\boldsymbol{Q}_k-\boldsymbol{\mu}_k)^t\boldsymbol{\Gamma}_k^{-1}(\boldsymbol{Q}_k-\boldsymbol{\mu}_k)\right] \tag{20}$$

The BIC value of each generated cluster is then compared with the joint BIC value of its split into two clusters, and the clusters associated with the smaller value are selected (the smaller BIC value the better fitting is). The BIC value for ω_{c_k} is computed according to the following equation:

$$\mathrm{BIC}(\omega_{c_k})=-2\log L_f\left[\hat{\boldsymbol{\Theta}}_k;\boldsymbol{Q}_k\in\omega_{c_k}\right]+2M\log n_k \tag{21}$$

where $\hat{\boldsymbol{\Theta}}_k=\left[\hat{\boldsymbol{\mu}}_k,\hat{\boldsymbol{\Gamma}}_k\right]$ is the maximum likelihood estimate of the M-dimensional normal distribution. $\boldsymbol{\mu}_k$ and $\boldsymbol{\Gamma}_k$ denotes the M-dimensional means vector and the $M\times M$ dimensional covariance matrix, respectively. n_k is number of pixels in $\boldsymbol{Q}_k$. $2M$ is the number of free parameters. The likelihood function L_f is built as $L_f(\cdot)=\prod f(\cdot)$.

The joint BIC of $\omega_{c_{k_1}}$ and $\omega_{c_{k_2}}$ can be also computed in a similar way:

$$\mathrm{BIC}(\omega_{c_{k_1}},\omega_{c_{k_2}})=-2\log L_f'\left[\hat{\boldsymbol{\Theta}}_k';\boldsymbol{Q}_k\in\omega_{c_k}\right]+4M\log n_k \tag{22}$$

where $\hat{\boldsymbol{\Theta}}_k'=\left[\hat{\boldsymbol{\Theta}}_{k_1},\hat{\boldsymbol{\Theta}}_{k_2}\right]$ is the maximum likelihood estimate of the considered two M-dimensional normal distributions, and $\hat{\boldsymbol{\Theta}}(k_1)$ and $\hat{\boldsymbol{\Theta}}(k_2)$ represent the individual distribution parameters for $\omega_{c_{k_1}}$ and $\omega_{c_{k_2}}$, respectively. In this case, the total number of parameters is $4M$. L'_f is the likelihood function from the joint probability density function.

An additional merging operation is applied if necessary to ensure that the final output number of clusters is within the defined range G [106]. After applying the x-means clustering, the final output includes: 1) the optimal number k' of major changes; 2) the detected major changes in Ω_c (i.e., the level L_1 of the hierarchical tree structure). The adopted x-means method for identifying the hidden change classes inside of Ω_c are described in TABLE 5-1. Note that BIC is just one of the choices for the optimal model selection. However, it is a reliable criterion especially for normal distributions. Other test criteria, such as Akaike Information Criterion (AIC) and Minimum Description Length (MDL) may also be used [107], [108].

TABLE 5-1 *X*-MEANS ALGORITHM FOR THE AUTOMATICAL CHANGE CLUSTERING.

Inputs: M-dimensional PCs of the data in Ω_c, a given range $G = [k_0, k_0+t]$

***Step* 1**: apply the conventional k-means with the initial $k=k_0$ to obtain k_0 classes: $\omega_{c_1},\ldots,\omega_{c_{k_0}}$;

***Step* 2**: apply the operations from ***Step* 3** to ***Step* 5** to all class ω_{c_k}, $k = 1, 2,\ldots, k_0$;

***Step* 3**: for a given class ω_{c_k}, divide it into two ($\omega_{c_{k_1}}$ and $\omega_{c_{k_2}}$) by k-means with k equal to 2;

***Step* 4**: calculate the BIC value for both $\mathrm{BIC}(\omega_{c_k})$ and $\mathrm{BIC}(\omega_{c_{k_1}},\omega_{c_{k_2}})$;

***Step* 5**: compare the BIC values of the above two models, and make the decision of division according to the following rules:

If $\mathrm{BIC}(\omega_{c_k}) > \mathrm{BIC}(\omega_{c_{k_1}},\omega_{c_{k_2}})$, the division is continued. Let $\omega_{c_k} \leftarrow \omega_{c_{k_1}}$, push the data related to $\omega_{c_{k_2}}$ onto the stack. Return to ***Step* 3**;

If $BIC(\omega_{c_k}) \le BIC(\omega_{c_{k_1}},\omega_{c_{k_2}})$, there is no division for class ω_{c_k}. Extract the stacked data, return to ***Step* 3**. If the stack is empty, then move to ***Step* 6**.

***Step* 6**: update all the classes and stop the iteration, thus temporary k_t classes are obtained;

***Step* 7**: make decision for merging operation:

If $k_t \le k_0+t$, move to the ***Step* 9**;

If $k_t > k_0+t$, move to the next merging step.

***Step* 8**: sort all the ω_{c_k} according to the amount of data in an ascending order, update the subscript as α and β ($\alpha,\beta = 1,\ldots, k_t$).

while $k_t > k_0+t$,

Apply the BIC as a round robin to ω_{c_α} and merge ω_{c_α} and ω_{c_β}, where $\alpha \neq \beta$ and $\beta > \alpha$.

If $\mathrm{BIC}(\omega_{c_\alpha}) > \mathrm{BIC}(\omega_{c_\alpha},\omega_{c_\beta})$, merge ω_{c_α} into ω_{c_β}

Else do not merge.

Update k_t

(Note that any ω_{c_α} and ω_{c_β} is restricted to be used only one time.)

***Step* 9**: update all the classes with $k'=k_t$.

Outputs: 1) the optimal number of classes k'; 2) clustering results, in which each pixel within the input PCs of Ω_c belongs to a specific class ω_{c_k}, $k = 1, 2,\ldots, k'$.

To define a reliable range G for the clustering process, the initial number of classes should be identified, which is the lower bound k_0 ($k_0 \geq 2$) in the x-means. k_0 should be small enough to include the minimum number of change classes that can be directly recognized. To perform a reliable choice of this parameter, we applied a method based on the analysis of the compressed change direction representation proposed in [47]. Instead of directly computing the angular distance in the original feature space, we computed it on the selected M-dimensional PCs of Ω_c as follows:

$$\varphi(\boldsymbol{Q}(i,j)) = \arccos\left(\frac{1}{\sqrt{M}} \left(\sum_{m=1}^{M} Q^m(i,j) \Big/ \sqrt{\sum_{m=1}^{M} (Q^m(i,j))^2} \right) \right) \tag{23}$$

where $\varphi(\boldsymbol{Q}(i,j))$ is the compressed change direction of $\boldsymbol{Q}(i,j)$, and $Q^m(i,j)$ is the m-th component of vector $\boldsymbol{Q}(i,j)$. In this way we emphasize in the direction variable only the possible changes associated with Ω_c. The first PCs can properly model the changes that we are looking for. Thus the modes of the obtained distribution on the compressed change direction $\varphi(\boldsymbol{Q}(i,j))$ can be recognized as the initial number k_0 of major changes existing in Ω_c (see Fig.5-4). The upper bound of the range G is defined by adding a small integer value t to k_0. t is in the order of few units and takes into account the intrinsic uncertainty of defining k_0 by analyzing $\varphi(\boldsymbol{Q}(i,j))$.

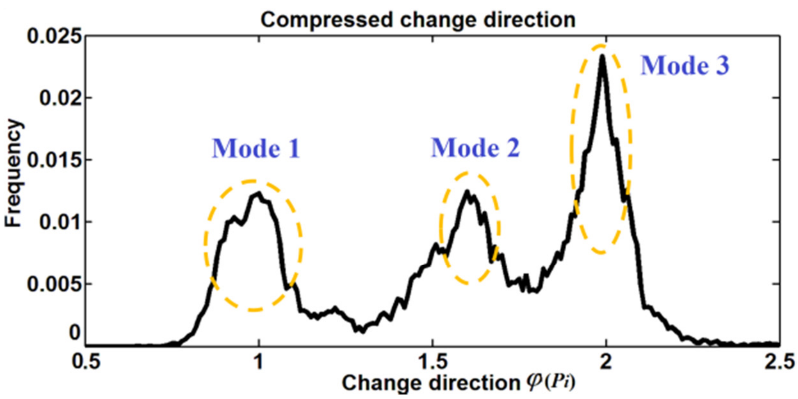


Fig.5-4 Example of definition of the initial cluster number (k_0) based on analyzing of the compressed change direction.

Once the major change classes in Ω_c have been recognized and separated by using the adopted clustering algorithm, the root node splits into different child nodes at L_1 in the tree. Each node corresponds to one major change class (i.e., $\omega_{C_1}, \omega_{C_2}, \ldots$). For each major change $\omega_{C_1}, \omega_{C_2}, \ldots$ the spectral homogeneity of SCVs is tested according to (4). As an example let us consider the first child node associated to class ω_{C_1}. The SAD of ω_{C_1} is computed as $\vartheta(\boldsymbol{x}(i,j), \boldsymbol{S}_{\omega_{c_1}})$ for each $\boldsymbol{x}(i,j) \in \omega_{C_1}$. If for a given node convergence is not reached (e.g., in our example it means that the standard deviation of $\vartheta(\boldsymbol{x}(i,j), \boldsymbol{S}_{\omega_{c_1}})$ is larger than a given threshold) then all the above operations (i.e., PCA, x-means, stop criterion evaluation) are iterated by considering only the SCVs of pixels $\boldsymbol{x}_i$ in the considered node (e.g., ω_{C_1} in our example). Once all the nodes at L_1 are processed, the algorithm moves to the next level. The hierarchical decomposition is applied to each node in every level of the tree until the convergence is reached for all of them (see Fig.5-5). This happens when all the nodes satisfy the homogeneous condition in (4). The last node of each branch is a leaf node and corresponds to one change endmember in $\Omega_e=\{e_1, e_2,\ldots,e_K\}$. Note that at convergence change endmembers can appear at different levels of the tree. The block scheme of this step is shown in Fig.5-5.

5.2.3 Generation of the Change-Detection Map by Endmember Clusters Merging

After identifying K change endmembers $\Omega_e =\{e_1, e_2,\ldots,e_K\}$, the pixels in the uncertain class Ω_u derived in the first step of pseudo-binary CD are considered. These pixels are assigned to one of the change endmembers or to the no-change class on the basis of spectral similarity. SAD (see (4)) is computed between the SCV $\boldsymbol{x}(i,j)$ ($\boldsymbol{x}(i,j) \in \Omega_u$) and the reference spectra $\boldsymbol{S}_{e_\varepsilon}$ (i.e., the average spectrum of each detected change endmember in Ω_e and of the no-change endmember e_n). Then $\boldsymbol{x}(i,j)$ is assigned to the class with the minimum distance value, i.e.,

$$\boldsymbol{x}(i,j) \in \underset{e_\varepsilon \in \{\Omega_e, e_n\}}{\arg\min} \left\{ \vartheta(\boldsymbol{x}(i,j), \boldsymbol{S}_{e_\varepsilon}) \right\} \tag{24}$$

where $\vartheta(\boldsymbol{x}(i,j), \boldsymbol{S}_{e_\varepsilon})$ denotes the SAD distance between $\boldsymbol{x}(i,j)$ ($\boldsymbol{x}(i,j) \in \Omega_u$) and a given reference spectrum $\boldsymbol{S}_{e_\varepsilon}$. The final CD map is generated by merging the results obtained in the three sets of changed, uncertain and unchanged pixels (see in Fig.5-1).

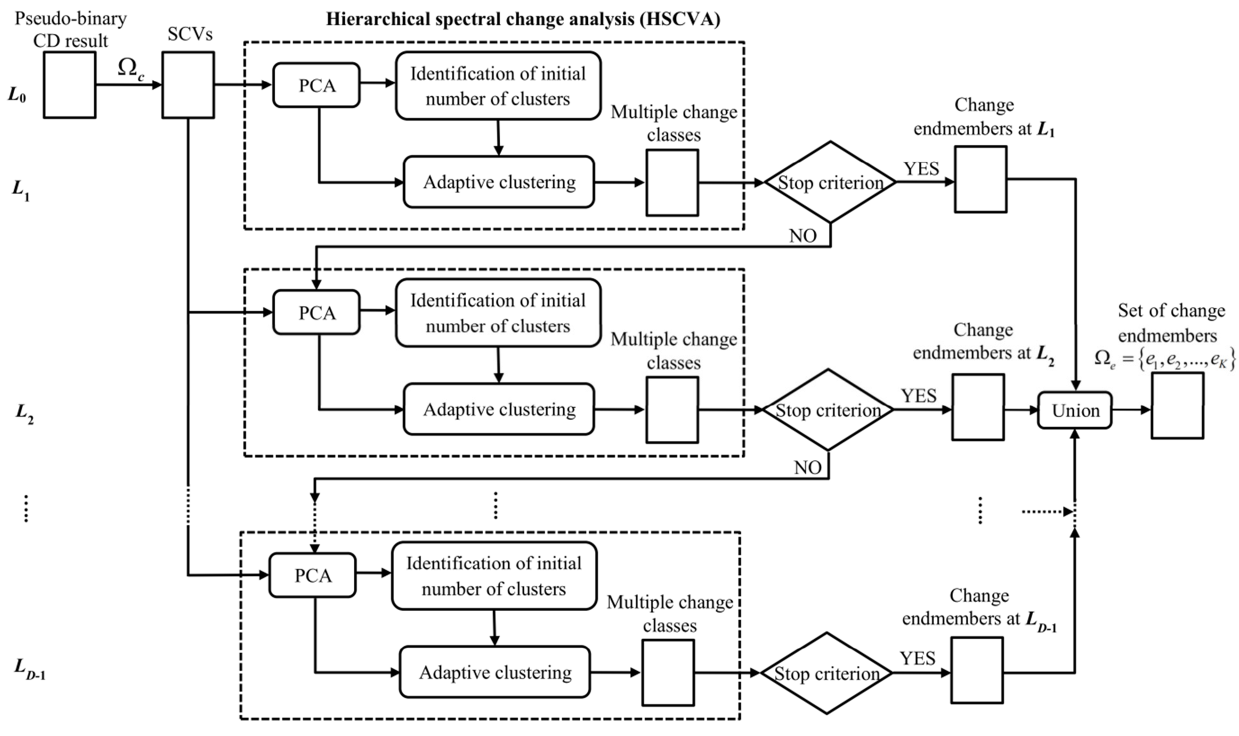


Fig.5-5 Block scheme of the HSCVA step in the proposed change-detection approach.

5.3 Data Set Description and Design of Experiments

5.3.1 Description of Data Sets

Data Set 1: Simulated Hyperspectral Camera Data Set

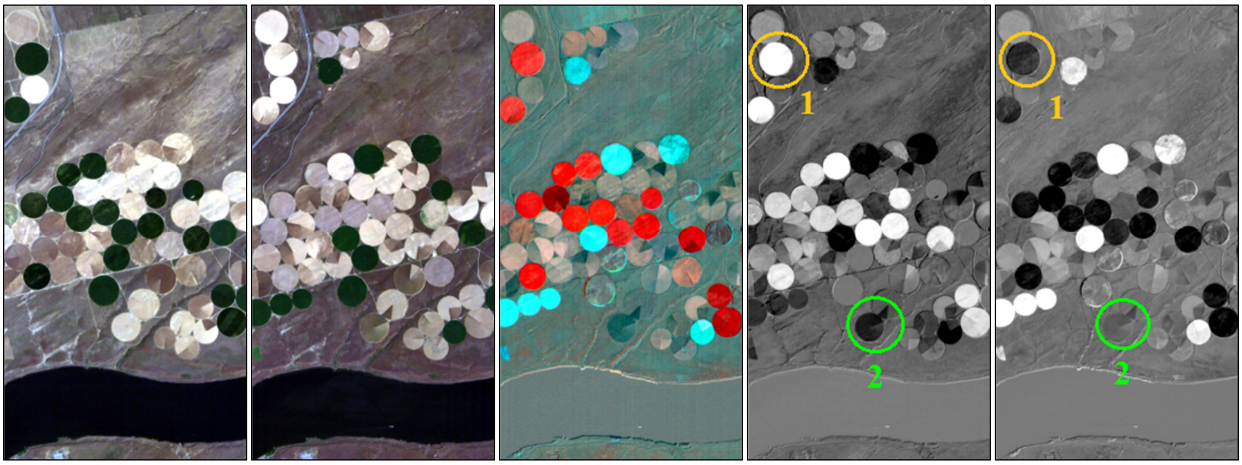
The first data set is the same as introduced in Chapter 4, Section 4.4.1 that acquired by a commercial hyperspectral camera (Nuance FX, CRI Inc.) [101] (see Fig.5-6.a). Ten change classes were simulated. Note that in this case, the same simulation setup was conducted three times by varying the position of tiles, thus generating three simulated multitemporal datasets. Each one is composed of $\boldsymbol{X}_1$ and one among the three simulated $\boldsymbol{X}_2$. Fig.5-6 (b) shows one of the simulated images, and Fig.5-6 (c) presents the corresponding change reference map, which includes ten change endmembers. The performance indices for this data will be presented as the average values over the three simulated data sets.

(a) (b) (c)

Fig.5-6 False color composite (R: 710nm; G: 620nm; B: 510nm) of (a) the hyperspectral image acquired by the Nuance FX hyperspectral camera ($\boldsymbol{X}_1$); (b) one of the simulated image ($\boldsymbol{X}_2$) with changes; (c) Reference map (ten changes in different colors, no-change class in white color).

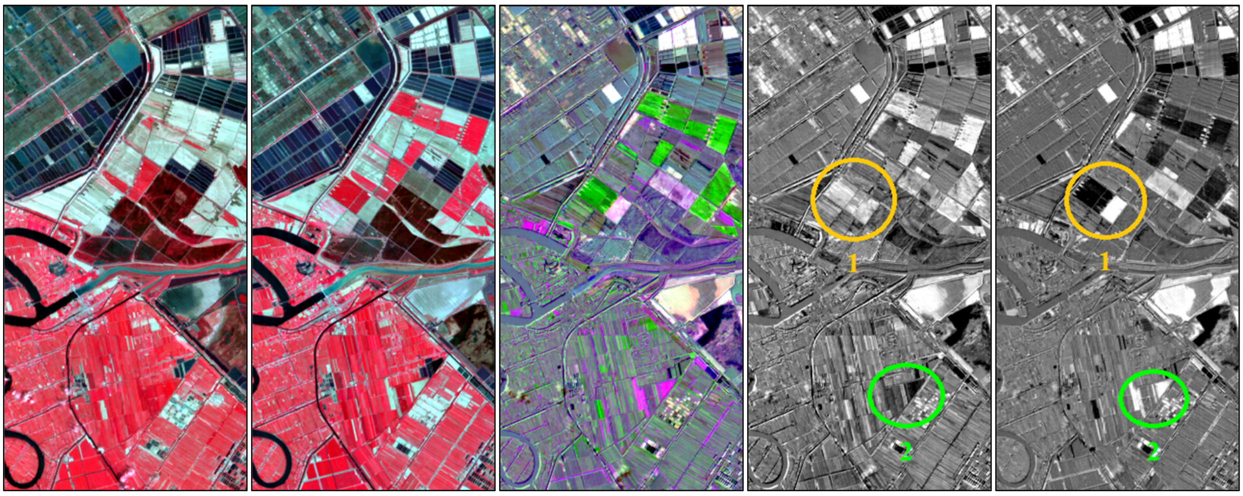
Data Set 2: Hyperion satellite images of an irrigated agricultural area

The second data set is the same as introduced in Chapter 4, Section 4.4.3 that acquired by the Hyperion sensor mounted onboard the EO-1 satellite on May 1st, 2004 (i.e., $\boldsymbol{X}_1$, see Fig.5-7.a) and May 8th, 2007 (i.e., $\boldsymbol{X}_2$, see Fig.5-7.b). The study area covers an irrigated agricultural land of Hermiston city in Umatilla County, Oregon, United States. Land-cover changes include the transitions among the crops, soil, water and other land-cover types. The changes occurred in the crop land are mainly due to the vegetation water content that affected the irrigation condition in the field (see the circles on the image, which correspond to the radius of the irrigation system), and to the difference of the crop growth situation. Fig.5-7 (c) represents a false color composite of spectral channels in $\boldsymbol{X}_D$. Different colors indicate possible kinds of change classes, whereas gray areas represent the unchanged pixels. The same change class can be described differently in different wavelengths (e.g., see Fig.5-7 (d) and (e) where the same kind of change is highlighted in orange and green circles and has different behaviors in bands 30 and 40 of $\boldsymbol{X}_D$). Accordingly the two examples given in Fig.5-7 do not fully describe the complexity of the problem.



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

Fig.5-7 Hyperion images acquired on an irrigated agricultural area. False color composite (R: 650.67nm, G: 548.92nm, B: 447.17nm) of the original images acquired in (a) 2004 ($\boldsymbol{X}_1$) and (b) 2007 ($\boldsymbol{X}_2$); (c) composite three SCVs channels (R: 1729.70nm, G: 1023.40nm, B: 752.43nm); single SCVs channel of (d) band 30 (650.67nm) and (e) band 40 (752.43nm).



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

Fig.5-8 Hyperion images acquired on a wetland area in China. False color composite (R: 752.43nm, G: 650.67nm, B: 548.92nm) of the original images acquired in (a) 2006 ($\boldsymbol{X}_1$) and (b) 2007 ($\boldsymbol{X}_2$); (c) composite of SCVs channels (RGB: 1729.70nm, 752.43nm, 650.67nm); selected SCVs channels: (d) band 52 (874.52nm) and (e) band 158 (1729.70nm).

Data Set 3: Hyperion images of wetland agricultural area

Another pair of bi-temporal Hyperion hyperspectral images with a size of 252×526 pixels, acquired on May 3, 2006 ($\boldsymbol{X}_1$) and April 23, 2007 ($\boldsymbol{X}_2$) in a wetland agricultural area in Yancheng, Jiangsu province, China, was used in the experiments. After applying the same pre-processing used for the previous data set, 132 bands were selected: 13-53, 83-96, 101-118, 135-164, 188-199 and 202-218. Also for this dataset we do not have any available ground truth. False color composite images of the bi-temporal data are shown in Fig.5-8 (a) and (b). In this scenario, the land-cover classes mainly include the agricultural cropland, seafood farm ponds, and offshore shoals vegetation (e.g., spartina alterniflora, suaeda and reed). During the study period, the actual land-cover changes included transitions among vegetation (most were in the crop field), water area, seafood farm ponds and some buildings. Fig.5-8 (c) shows a false color composite image of $\boldsymbol{X}_D$, and Fig.5-8 (d) and (e) present some selected channels of $\boldsymbol{X}_D$. Similarly as before, the false color composition does not fully describe the complex CD problem. However, it gives an idea about where the changes occurred.

5.3.2 Design of Experiments

The proposed CD approach has been applied to the three hyperspectral data sets. For the synthetic bi-temporal hyperspectral images, the same procedure was conducted on three simulated data sets. In this case, the first step of pseudo-binary CD was neglected as the general change class Ω_c is explicitly defined by the change simulation step. Thus we directly focused on the pixels in Ω_c and tried to identify different change endmembers inside it. Performance is assessed quantitatively on the three reference maps. The final performance indices are given as the average accuracy over the three simulated datasets. For the two Hyperion hyperspectral remote sensing data sets, the proposed method was applied starting from the pseudo-binary CD step and the three clusters (Ω_c, Ω_u and ω_n) where generated. The value δ was set such that the Ω_u class includes 25% of the pixels in ω_n. After obtaining the general change class Ω_c, T_σ was set to drive the decomposition of the root node and to build the hierarchical tree for change endmember detection. T_σ is a user dependent parameter and controls the level of spectral homogeneity of the detected change endmembers. The smaller the threshold value T_σ, the higher the homogeneity level is and thus the number of change endmembers, and viceversa. In practical applications, the threshold should be selected taking into account the desired sensitivity to subtle changes. In our experiments trials were carried out with different values of T_σ, achieving different trade-offs in terms of endmember homogeneity.

After the initialization of Ω_c (i.e., root node of the tree), the identification of multiple change endmembers was done by using the proposed HSCVA step. The initial number of k_0 was defined based on the compressed change direction method described in Section 5.2.2, and t was set equal to 3 to define the upper bound of U. The final CD map was obtained when all change endmembers were generated and the pixels in Ω_u were assigned to one of them or to the unchanged endmember. The results obtained by the proposed method were compared with the ones obtained by the popular unsupervised k-means and fuzzy C-means (FCM) clustering methods. The two reference methods were applied to the subset of PCs selected by the proposed method for the root node, i.e., the ones that contain most of the information for Ω_c. The class number k of k-means and FCM was fixed on the basis of the proposed method outcome. In this way, we give clear advantage to the reference techniques that have not the intrinsic capability to estimate the number of expected change endmembers. This choice implicitly penalizes the proposed method. To reduce the uncertainty due to the random initialization in the reference methods, we ran them 200 times. The final accuracy was calculated as the average over 200 trials.

To evaluate the CD results both quantitative and qualitative assessments were carried out for each of the three considered datasets. For the synthetic data set, the quantitative assessment was based on the CD accuracy (i.e., endmember accuracy and kappa accuracy) and error indices obtained according to the reference maps. In addition, the average endmember distance has been computed to assess the average endmember separability. To this end, pair-wise Bhattacharyya distance was computed among all the pairs of change endmembers. For two generic detected change endmembers e_α and e_β (α, $\beta \in [1, K]$ and $\alpha \neq \beta$), the Bhattacharyya distance $B_{\alpha,\beta}$ is calculated as follows:

$$B_{\alpha,\beta} = \frac{1}{8}(\boldsymbol{\mu}_\alpha - \boldsymbol{\mu}_\beta)^T \left\{ \frac{\boldsymbol{\Gamma}_\alpha + \boldsymbol{\Gamma}_\beta}{2} \right\}^{-1} (\boldsymbol{\mu}_\alpha - \boldsymbol{\mu}_\beta) + \frac{1}{2} \ln \left\{ \frac{\left| (\boldsymbol{\Gamma}_\alpha + \boldsymbol{\Gamma}_\beta)/2 \right|}{\left|\boldsymbol{\Gamma}_\alpha\right|^{1/2} \left|\boldsymbol{\Gamma}_\beta\right|^{1/2}} \right\} \tag{25}$$

where $\boldsymbol{\mu}_\alpha$ and $\boldsymbol{\mu}_\beta$ denote the mean vectors, $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ represent the covariance matrices of change endmembers α and β, respectively. The higher distance the better the class separability, and viceversa. The average pairwise Bhattacharyya distance computed on all pairs of change endmembers gives indication of the overall class separability. In the following we will refer to it as multi-class Bhattacharyya distance.

The CD results were also analyzed qualitatively by comparing: 1) the obtained CD maps; 2) the 2-D scatterplots of change endmembers in the feature space (i.e., the first PC versus the second PC on Ω_c); 3) the spectral signatures of all the detected change endmembers in $\boldsymbol{X}_D$ with the ones obtained with reference techniques.

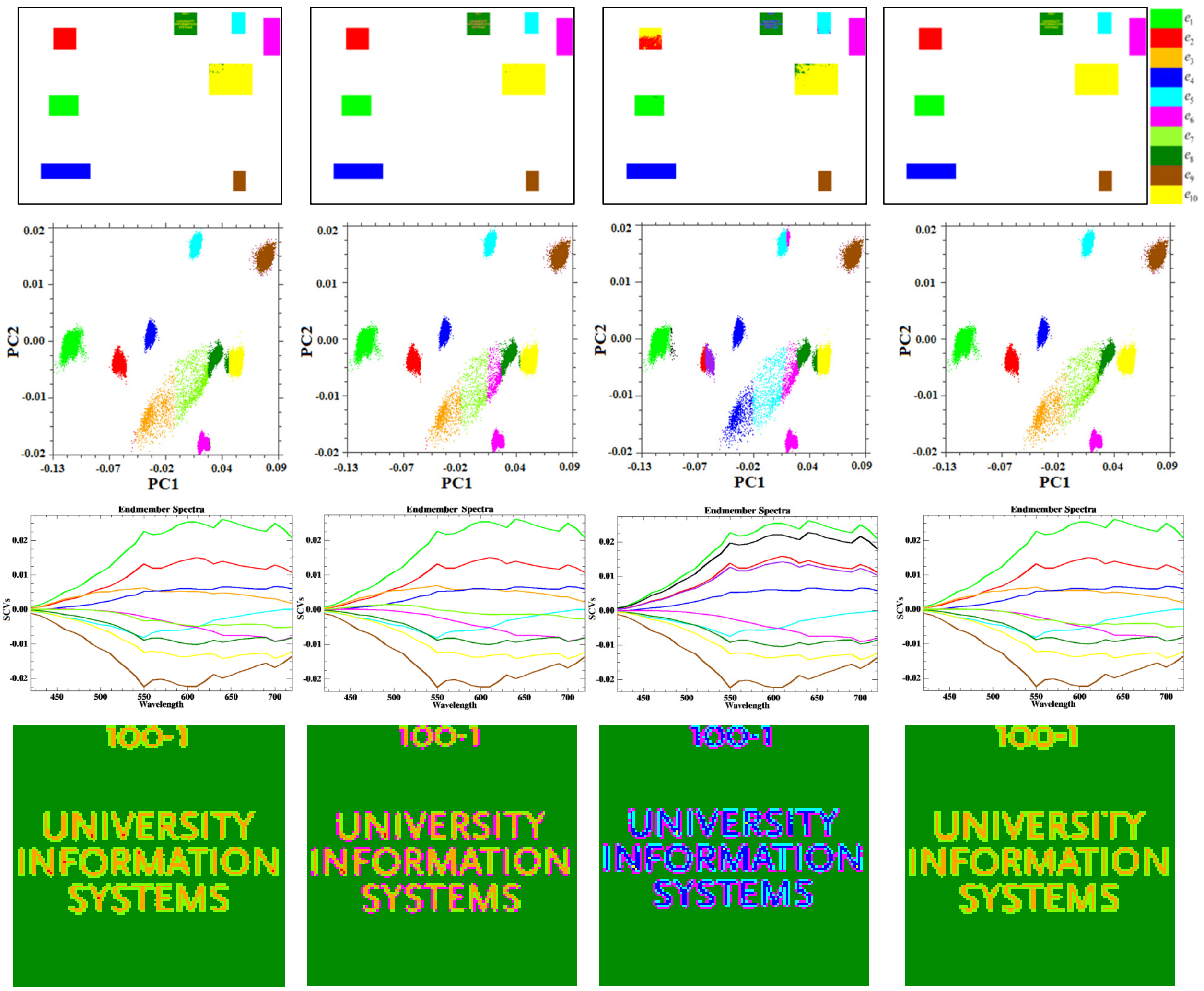
5.4 Experimental Results

5.4.1 Simulated Hyperspectral Data Set

For the simulated data, experimental results were obtained by fixing the value of T_σ to 0.05 for all the three image pairs. The average kappa accuracy (κ) and the average multi-class Bhattacharyya distance obtained by the three considered methods are shown in TABLE 5-2. As one can see, the proposed method obtained both the highest kappa accuracy and the highest average Bhattacharyya distance.

TABLE 5-2 AVERAGE KAPPA ACCURACY AND MULTI-CLASS BHATTACHARYYA DISTANCE OBTAINED BY THE THREE CONSIDERED METHODS ON THE SIMULATED DATA SETS.

Method	Average κ	Average multi-class *Bhattacharyya* distance
PCA *k*-means	0.9772±0.0007	5.28
PCA FCM	0.9002±0.0012	5.03
Proposed method	0.9930±0.0009	5.91



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

Fig.5-9 CD results obtained on the simulated hyperspectral data set. Results provided by: (a) the proposed method, (b) k-means, and (c) FCM, (d) ground truth. From up to down, each row represents: (1) CD maps (or reference map); (2) 2-D scatterplots of change classes in the feature space; (3) SCV signatures of detected changes; (4) a subset from results in (1).

Let us now analyze one of the three simulated cases in greater detail (see Section 5.3.1, Fig.5-6 .b). In this case, the complete tree has a structure with 3 levels and 14 nodes, where 10 of them are leaf nodes identified as change endmembers. The CD maps obtained by the proposed method and the reference ones are shown in the first row of Fig.5-9. Fig.5-9 (a)-(c) reports the results of the proposed method, the reference k-means and FCM, respectively. Fig.5-9 (d) shows the reference map. Each color corresponds to a specific detected change endmember, whereas the unchanged pixels are in white. In the second row, 2-D scatterplots of the detected change classes are shown in the feature space of first two PCs extracted from pixels in Ω_c. The spectral behaviors of the change endmembers in the SCV domain are presented in row 3. Tiles extracted from the whole CD maps are illustrated and further compared in row 4. Accuracies and error indices obtained according to the reference data are summarized in TABLE 5-3.

TABLE 5-3 KAPPA ACCURACY, NUMBER OF DETECTION ERRORS AND MULTI-CLASS BHATTACHARYYA DISTANCE OBTAINED BY THREE CONSIDERED METHODS ON ONE OF THE SIMULATED DATA SETS.

Method	Endmember accuracy (%)										κ	Tot. errors (pixel)	Multi-class *Bhattacharyya* distance
	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}			
PCA *k*-means	100.00	99.97	88.56	100.00	99.99	91.37	39.34	98.15	100.00	97.42	0.9770	1367	5.49
PCA FCM	99.77	57.30	0.00	100.00	97.10	97.20	0.00	97.92	100.00	94.30	0.9007	2218	4.93
Proposed method	100.00	99.93	92.10	100.00	100.00	99.94	86.60	99.46	100.00	99.07	0.9933	650	6.22

As we can see from Fig.5-9, the proposed method detected the expected changes on this simulated data set accurately. In particular, it identified properly the change classes in a hierarchical way, and it was not affected by the problem on minority classes. The subtle changes with small amount of pixels (e.g., change of letters and their edges) were also detected in a precise way (see row 4 in Fig.5-9). On the contrary, despite the conventional *k*-means and FCM received as input the true number of change endmembers, their results were less accurate. This demonstrates the advantages of using the hierarchical analysis structure. A visual comparison of scatterplots confirms the better results produced by the proposed method with respect to the other techniques. The two reference methods obtained in overall good performances, but showed a higher error rate for some change endmembers (e.g., e_6 is confused with e_7 in *k*-means; and e_3 with e_4 in FCM). By comparing the SCV signatures of changes detected by the three methods (Fig.5-9 row 3 (a)-(c)) with the one of reference change map (Fig.5-9 row 3 (d)), we can observe: 1) higher similarity between results of the proposed method and the reference spectra; and 2) different kinds of change (i.e., change endmembers) have discriminable spectral behaviors in the SCV domain (see row 3 (a) in Fig.5-9), thus indicating the effectiveness of the proposed method in separating change information. The reference techniques detected some wrong change endmembers. For example, in the result of the FCM there are two couples of change endmembers with very similar spectral signatures. The first couple is represented by red and purple signatures, and the second is given by green and sienna signatures in Fig.5-9 (c) row 3. These changes were wrongly detected by the FCM method even by fixing the correct number of input classes.

The above analysis is confirmed by the numerical results in TABLE 5-3. We can observe that: 1) the proposed hierarchical method outperformed reference approaches in terms of Kappa accuracy and number of errors. The Kappa accuracy is the highest among the three (i.e., 0.9933 compared to 0.9770 for *k*-means and 0.9007 for FCM). The total error of the proposed method (i.e., 650 pixels) is significantly smaller than the ones of reference methods (i.e., 1367 pixels for *k*-means and 2218 pixels for FCM); 2) on each single change endmember, the two reference approaches resulted in significant errors (both omission and commission), whereas the proposed method exhibits the highest accuracy. This further confirms the difficulty of the reference methods to directly identify endmembers; 3) the multi-class Bhattacharyya distance values indicate that the proposed approach achieves the highest class separability (i.e., 6.22) with respect to the two clustering methods (i.e., 5.49 in *k*-means and 4.93 in FCM, respectively).

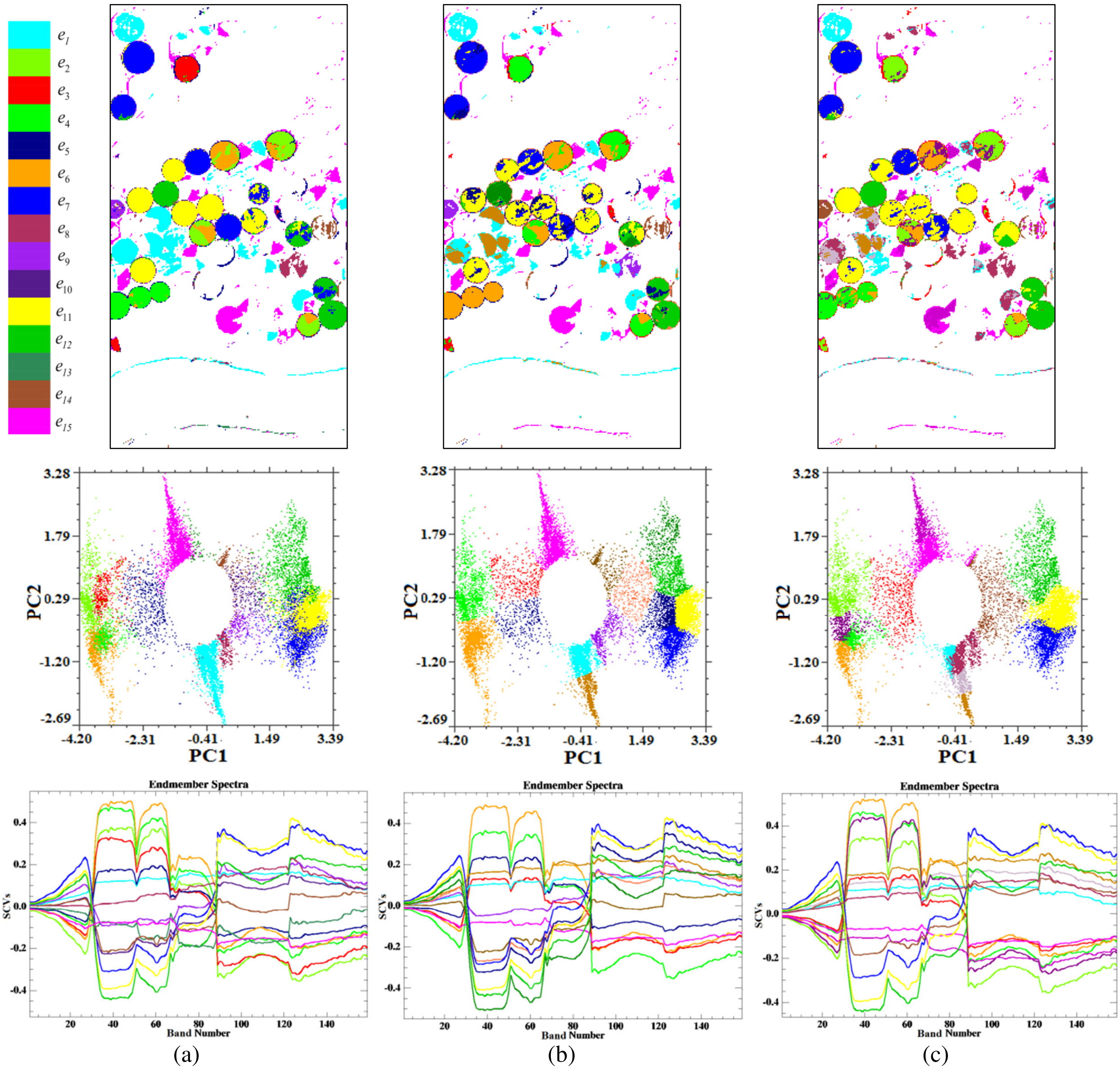


Fig.5-10 CD results obtained on the real Hyperion hyperspectral images on an agricultural area. Results provided by (a) the proposed method, (b) the *k*-means, and (c) the FCM. From row 1 to row 3: (1) change-detection maps; (2) 2-D scatterplots of all change classes in the feature space by using the first two PCs computed on pixels in Ω_c; (3) spectra of the detected changes in the SCV domain. The legend only applies to the proposed method results.

5.4.2 Hyperion Satellite Images of an Irrigated Agricultural Area

In this case the threshold T_σ was set to 0.13. The proposed method detected 15 change endmembers as leaf nodes in the hierarchical tree, which includes 4 levels and 20 nodes. Fig.5-10 illustrates CD results obtained by (a) the proposed hierarchical method, (b) the *k*-means, and (c) the FCM. From row 1 to row 3 the figure shows the CD maps, the 2-D scatterplots in the two-dimensional feature space (i.e., the first two PCs extracted from pixels in Ω_c), and the SCV signatures of all the detected changes, respectively. For the proposed hierarchical approach the 15 change endmembers are represented with different colors, whereas the no-change pixels are in white. For the two reference methods, the change clusters are also shown in different colors, but it is not possible to establish a direct correspondence among the legend given for the proposed method in Fig.5-10, and the colors used for the reference methods. Also in this case the number of clusters for the *k*-means and the FCM was fixed on the basis of the result produced by the proposed technique.

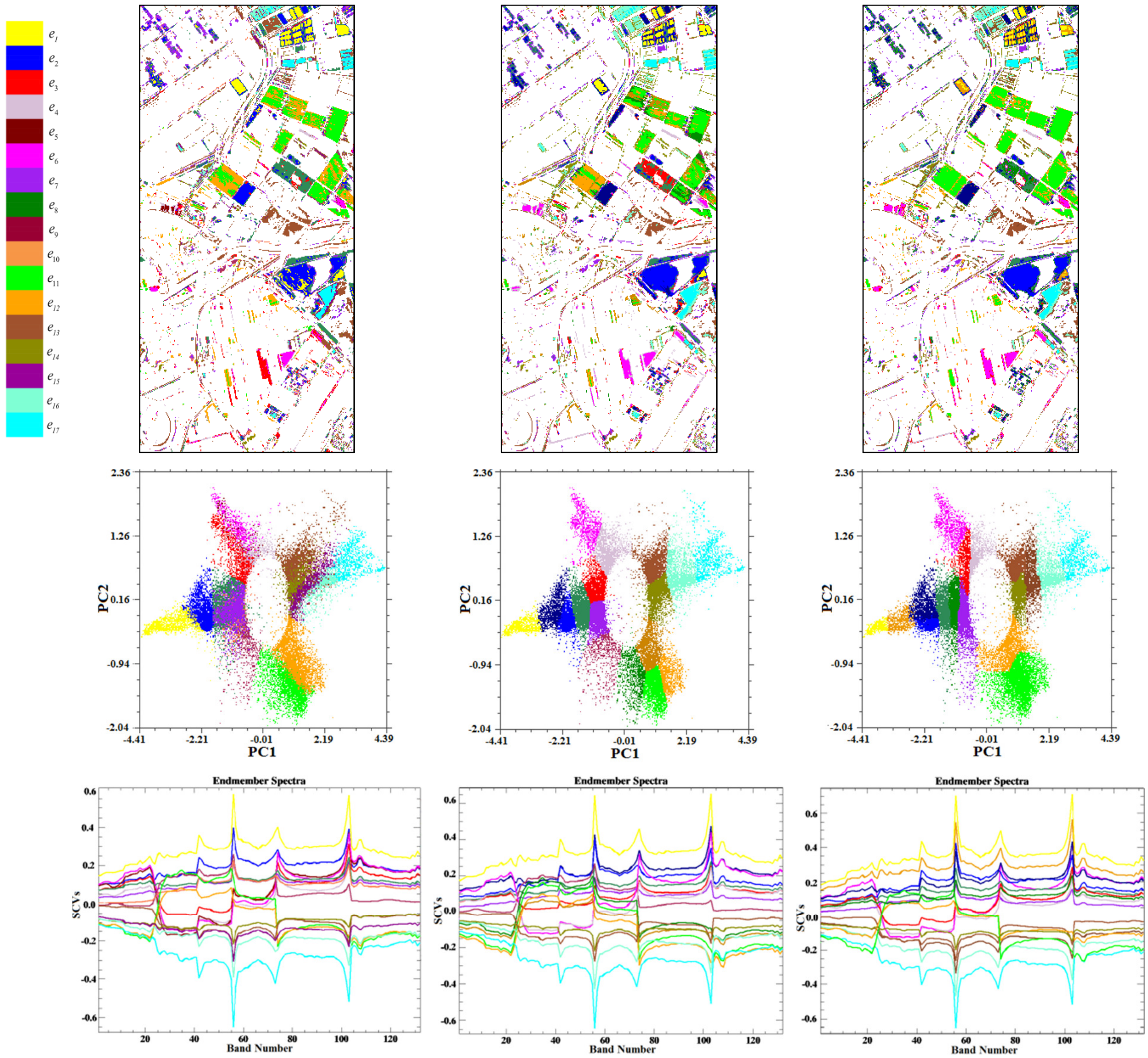
The proposed hierarchical CD approach obtained satisfactory results detecting change endmembers (validated by the detailed photointerpretation) and separating them according to the defined spectral homogeneity level. In greater detail, we can observe that: 1) The proposed method detected change endmembers according to the hierarchical analysis. On the contrary, the other two reference methods (that identify all the changes in a single step) ignore the intrinsic hierarchy of the change information in hyperspectral images. This increased the change detection errors (see also Fig.5-10 row 1, where the proposed method detects changes with a higher homogeneity than the two reference methods). 2) All the considered methods are able to discriminate multiple changes, but with different performance on the change separability of change endmembers. The multi-class Bhattacharyya distance values were 4.12 (proposed method), 3.78 (*k*-means) and 3.65 (FCM). The proposed method obtained the highest separability among all the detected change endmembers. 3) The generated spectra of change endmembers point out the differences of the SCV signatures, which illustrate the change separability of the different methods.

5.4.3 Hyperion Images of a Wetland Agricultural Area

On the third data set we carried out the same experiments as for the previous one. The threshold T_σ was set to 0.15. The hierarchical tree structure consisted of 5 levels with 27 nodes, where 17 change endmembers were detected according to 17 leaf nodes. As we can see from the CD results, in this case the proposed method also obtained satisfactory results. A qualitative analysis points out that the change endmembers were properly detected (see Fig.5-11). The multi-class Bhattacharyya distances for the three methods were 3.89 (proposed method), 3.49 (*k*-means) and 3.30 (FCM). Also in this case, the proposed hierarchical method achieved the highest multi-class separability, whereas the *k*-means and FCM resulted in a lower separability, despite the two reference methods are driven by the number of endmembers automatically detected by the proposed method.

5.5 Discussion and Conclusions

This chapter analyzed and discussed the change-detection problem in multitemporal hyperspectral images. A novel hierarchical spectral change analysis approach has been proposed to detect and identify multiple-change information in a partially-unsupervised way. Accordingly, the change endmembers are detected hierarchically by analyzing the spectral properties in the spectral difference domain $\boldsymbol{X}_D$. Moreover, the proposed hierarchical analysis can identify the discriminable spectral change endmembers from coarse to fine level leading to a better model, whereas the reference methods are based on a single step of processing only. Since in the CD-HS case, the number of change endmembers is usually high, those methods are generally not able to correctly identify all of them. Satisfactory results obtained on both the simulated and real multitemporal hyperspectral remote sensing images confirmed the effectiveness of the proposed CD method.



(a) (b) (c)

Fig.5-11 CD results obtained on the real Hyperion images on a coastal wetland agricultural area in China. Results provided by (a) the proposed method, (b) the *k*-means, and (c) the FCM. From row 1 to row 3: (1) CD maps; (2) 2-D scatterplots of all change classes in feature space by using first two PCs computed on pixels in Ω_c; (3) SCV spectra of the detected changes (17 changes in different colors, no-change class in white). The legend only applies to the proposed method results.

The main contributions of this chapter are as follows: 1) proposal of a technique for addressing the challenging CD-HS problem, by considering the difference of spectral change behaviors in the spectral difference domain $\boldsymbol{X}_D$ at different spectral detail scales; and 2) proposal of an approach that models the detection of multiple changes in a hierarchical way, to identify the change information and separate different kinds of changes (major change, subtle change and finally change endmembers) according to their spectral difference. In this way, we progressively decompose the complex problem into several specific sub-problems, focusing on each single portion of the multiple-change information. This makes it possible to discover the difference among similar changes by decreasing the difficulty of detection. Moreover, the proposed approach is designed in a partially-unsupervised way, which just requires a manual initialization for the clustering, thus it fits most of actual applications, for which often the ground truth is not available.

A minor limitation of the proposed method consists in the use of CVA for the pseudo binary CD step. By computing the magnitude of SCVs, small portions of the change information might be lost after compression, thus causing missed alarms in the final CD map. Although a proper setting of margin δ may limit this problem, the high dimensionality of hyperspectral data may still produce errors. Another issue to consider is the tuning of the threshold value (i.e., T_σ), which impacts on the final number of the output change endmembers. T_σ should be fixed in order to tune the sensitivity of the method according to the end-user requirements. This can be done considering the fact that T_σ has a clear physical meaning with respect to the sensitivity of the method. Although additional investigations should be done to define a possible automatic technique for the detection of the optimal threshold, we point out that the selection of T_σ is more simple and reliable than the selection of the number of endmembers in standard clustering methods.

As future development of this work, the robustness of the proposed method will be tested on the available multitemporal hyperspectral images showing differences in illumination conditions and no real change. Moreover, we plan to: i) consider in the proposed technique also the spatial information in order to increase the robustness and the accuracy of the CD results; ii) define a reliable automatic technique for the detection of the above mentioned threshold; iii) define alternative methods for the identification of change endmembers; iv) investigate the CD-HS problem on data sets for which an exhaustive ground truth is available.

Chapter 6

A Novel Unsupervised Multitemporal Spectral Unmixing for Detecting Multiple Changes in Hyperspectral Images

This chapter presents a novel unsupervised multitemporal spectral unmixing (MSU) approach to address the challenging multiple-change detection problem in bi-temporal hyperspectral images. Differently from the state-of-the-art methods that are mainly designed at pixel level, the proposed technique investigates the spectral-temporal variations at subpixel level. The considered CD problem is analyzed in a multitemporal domain, where a bitemporal spectral mixture model is defined to analyse the spectral composition within a pixel. Distinct multitemporal endmembers (MT-EMs) are extracted according to an automatic technique. Then a change analysis strategy is designed to distinguish the change and no-change MT-EMs. An endmember grouping scheme is applied to the changed MT-EMs to detect the unique change classes. Finally, the considered multiple-change detection problem is solved by analysing the abundances of the change and no-change classes and their contribution to each pixel. The proposed approach has been validated on both simulated and real multitemporal hyperspectral datasets presenting multiple changes. Experimental results confirmed the effectiveness of the proposed method.

6.1 Introduction

By taking advantage of the detailed spectral information in hyperspectral images, subtle changes (which are not visible when employing multispectral images) associated to the land-cover transitions are expected to be detected. Thus it is important to develop effective CD techniques that fully exploit the fine spectral variations in hyperspectral images to address new applications. However, due to the intrinsic properties of hyperspectral data, this task is highly challenging [86]. Examples of such properties are the high-dimensionality of the feature space, the information redundancy, the noise, and the presence of many possible change classes.

Only few literature papers can be found focusing on the topic of multitemporal CD-HS and even less that deal with the detection of multiple changes [42], [43], [86], [87]. Despite the usefulness of these approaches, they are all developed based on the assumption that each pixel in the considered images contains only one kind of land-cover material (pure-pixel theory). Accordingly, the final CD result associates a pixel only with a single specific kind of land-cover transition (i.e., vegetation to water, soil to building, etc.). However, given the geometrical resolution of hyperspectral images, mixed pixels are a common phenomenon that occurs in most of the cases. This phenomenon consists in a mixture of the light scatter-

This chapter is under revision for *IEEE Transactions on Geoscience and Remote Sensing*. Title : "Unsupervised Multitemporal Spectral Unmixing for Detecting Multiple Changes in Hyperspectral Images". Authors: S.Liu, L.Bruzzone, F.Bovolo, P.Du.

ing of more than one distinct substance located in the area on the ground covered by one pixel [109], [110], [111]. This mixture usually is caused either by the limited spatial resolution of the sensors included different targets in a single pixel or by the combination of the distinct materials into a homogeneous mixture [109], [110]. To solve this mixture problem, *spectral unmixing* techniques were developed aiming to detect the materials (termed *endmembers*) in the mixed pixels and to estimate their corresponding fractions (termed *abundances*). Thus the hyperspectral unmixing is actually an inverse problem. Despite single image unmixing has been widely investigated in the literature [104], [109], [110], [111], multitemporal unmixing has not been considered extensively. Land-cover material transitions within a single pixel are almost ignored in the available CD methods. The impact is a higher number of CD errors due to the spectral sensitivity of the hyperspectral data and to the poor investigation of the subpixel level spectral variations that is typical of state-of-the-art pixel-level CD methods. Therefore, it is necessary to consider the spectral mixture nature in the CD-HS studies and to develop advanced techniques for detecting and analyzing the subpixel level spectral changes.

Multitemporal unmixing (MU) has been only partially investigated to address the endmembers variability issue in order to increase the representativity of the extracted endmembers, thus improving either the land-cover classification or specific change monitoring (e.g., cropland, invasive species, forest, etc.) [112], [113], [114]. For CD, Du et al. [73] proposed a linear mixture model for analyzing the endmembers and abundances estimated from each single time image to address only a binary CD problem. Recently, a subpixel level CD approach was developed to investigate the multiple composition evidence within pixels, thus to increase the binary CD accuracy. However, it was designed in a supervised framework under the assumption that endmember samples are available [115]. No literature work can be found that address the challenging multiple-change detection problem in multitemporal hyperspectral images from the spectral unmixing point of view with unsupervised techniques.

In this chapter, a novel CD approach is proposed that is suitable and effective for detecting multiple change classes in hyperspectral images through the analysis of multitemporal spectral mixtures. To this end, a novel *multitemporal spectral unmixing* (MSU) technique is proposed. The proposed technique considers the spectral signatures in the *multitemporal domain* (i.e., stacked feature space), and identifies the different "*multitemporal endmembers* (MT-EMs)" associated to the change and no-changed classes. To overcome challenging issues like the high spectral variability and the insensitivity to the small size change classes, a patch scheme is adopted. Distinct MT-EMs are extracted from each patch of the images at the local level. Then their abundances are estimated at global level. Change analysis and endmember grouping are conducted to find unique change classes, thus generating the final change-detection map based on the abundance combination. The proposed MSU approach is validated on both simulated and real bi-temporal hyperspectral images. The experimental results confirm the effectiveness of the proposed method in performing multiple-change detection in multitemporal hyperspectral images.

The rest of the chapter is organized as follows. Section 6.2 illustrates the proposed MSU method in details. Section 6.3 describes the hyperspectral data sets and introduced the design of the experiments. The experimental results are analyzed and discussed in Section 6.4. Section 6.5 draws the conclusion of this work.

6.2 Proposed Approach

The considered CD problem can be formalized as to assign to each pixel in $\boldsymbol{X}_S$ a class label in $\Omega = \{\Omega_c, \omega_n\}$, where Ω_c is the set of the K possible change classes $\Omega_c = \{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}\}$ and ω_n is the no-change class. To this aim, we propose a novel automatic and unsupervised multitemporal spectral unmixing (MSU) approach that is suitable to analyze the spectral signature mixture among the change and no-change MT-EMs in $\boldsymbol{X}_S$, and thus to detect the multiple-change classes. The overall architecture of the proposed CD approach is illustrated in Fig.6-1. It mainly consists of four steps: 1) multitemporal images stacking and image patch generation; 2) multitemporal spectral unmixing; 3) change analysis; 4) abundance combination and CD map generation. More details on each step are given in the following sub-sections.

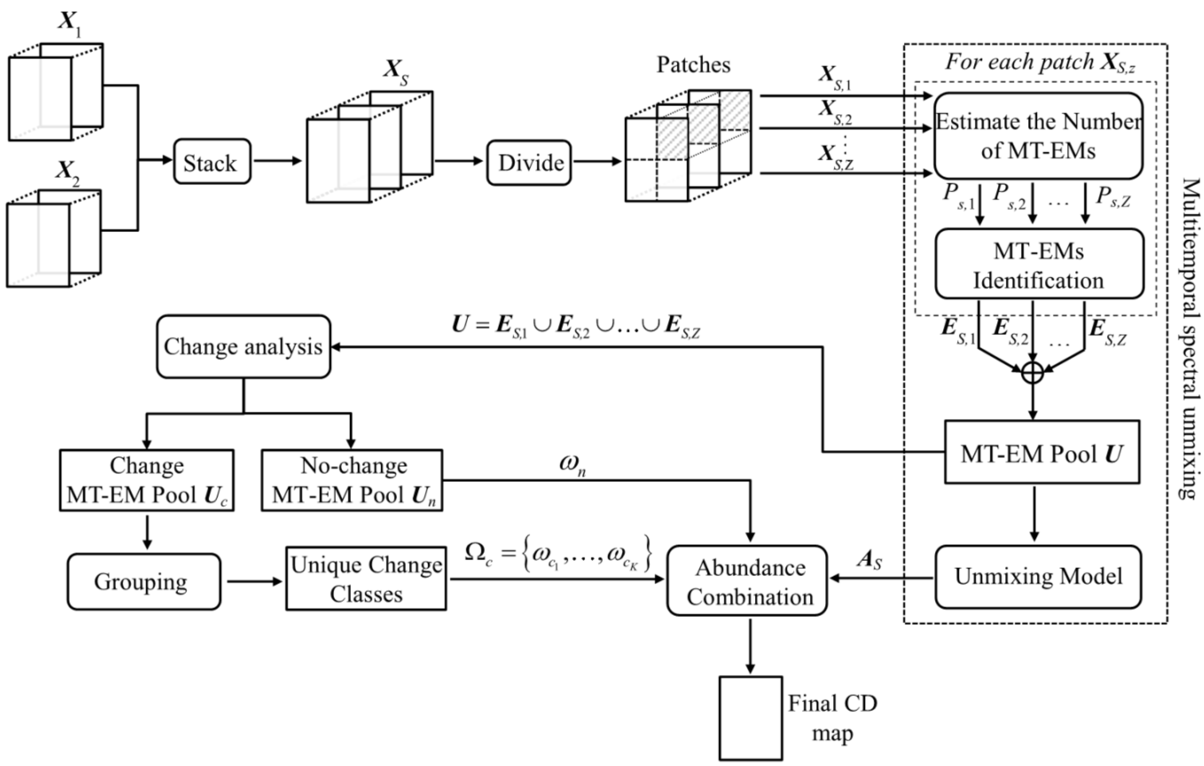


Fig.6-1 Architecture of the proposed CD approach based on multitemporal spectral unmixing.

6.2.1 Stacking of Multitemporal Images and Image-Patch Generation

In this step, the two B-Dimensional hyperspectral images $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are stacked into a $2B$-Dimensional image $\boldsymbol{X}_S$, whose spectral signatures (i.e., pixel vectors) include two components corresponding to $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively (see Fig.3-4). The next steps are all performed in the $\boldsymbol{X}_S$ domain. Then the $\boldsymbol{X}_S$ image is divided in a given number of sub-images called patches. Two reasons drive the use of a patch-based scheme. First of all, a higher number of MT-EMs can be detected in the stacked domain $\boldsymbol{X}_S$ than in the original single-date image domains (i.e., $\boldsymbol{X}_1$ or $\boldsymbol{X}_2$). Due to the fact that both change and no-change classes have discriminable spectral behaviors in $\boldsymbol{X}_S$, if P_1 and P_2 endmembers are identified in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively, theoretically a number of $P_s \leq P_1 \times P_2$ possible MT-EMs might be identified in $\boldsymbol{X}_S$. Therefore, it is very difficult to correctly identify all of them directly from the whole $\boldsymbol{X}_S$ feature space, especially for those small change classes defined only at a local level. The second reason is related to the endmembers variability in hyperspectral images [116], [117]. Due to both the land-cover spectral properties and variable external factors (e.g., atmospheric conditions, illumination, seasonal effects), spectral signatures that belong to the same material may vary within a hyperspectral image [117]. To overcome this drawback and identify discriminable endmembers in single hyperspectral images, several works have been presented in the literature that take into account the spectral variability. They can be mainly divided into two categories that consider endmembers either as sets (bundles) or as statistical distributions [117]. In this chapter, we use the "local endmembers" strategy introduced in [114], which belongs to the former group. It considers the spectral variability effects while identifying endmembers and performing unmixing. By following this idea, in our case the whole $\boldsymbol{X}_S$ image is divided into some regularly shaped subsets (patches) (see Fig.6-1). Let $\boldsymbol{X}_{S,z}$ be the z-th patch of $\boldsymbol{X}_S$ (z=1,…,Z), where Z is the defined number of patches. Note that Z is defined depending on the size of the image and the significance of the occurred change targets in the scene. Endmembers identification is performed on each patch. Due to the fact that the spectral classes are more uniform at the local level than at the global level, a large number of endmembers can be identified. Moreover, the patch-scheme better considers the small spectral variations and the material compositions generated in the local patches than a global-scheme.

6.2.2. Multitemporal Spectral Unmixing (MSU)

For each patch $\boldsymbol{X}_{S,z}$ (z =1,…, Z), MT-EMs are identified by one of the standard unmixing methods developed for single-date image. As an output we obtain: 1) The estimated number $P_{s,z}$ of the MT-EMs; 2) The identified distinct MT-EMs $\boldsymbol{E}_{S,z}$ according to $P_{s,z}$.

The number of MT-EMs $P_{s,z}$ is estimated automatically. Several algorithms can be found in the literature designed for this task. For example, we mention the Harsanyi-Farrand-Chang (HFC) algorithm [118], the noise whitened version of HFC (NWHFC) [119], the Hyperspectral Signal Identification by Minimum Error (HySIME) method [120], and the recently proposed Eigenvalue Likelihood Maximization (ELM) algorithm [121]. Note that any endmember number estimation and endmember extraction algorithms can be used within the proposed framework. The selection of an effective algorithm can ensure the quality of the extraction result, which depends on the considered data scenario and the specific applications. In our analysis, after several empirical trials, the ELM algorithm is selected as it proved to be accurate and stable. ELM is designed based on an empirical observation of the distribution of the differences of the eigenvalues from the correlation and the covariance matrices [121], which is totally parameter free and easy to be implemented. Then the popular Vertex Component Analysis (VCA) method [122] is used for extracting the $P_{s,z}$ MT-EMs $\boldsymbol{E}_{S,z}$. VCA is selected as the endmembers are the vertices of a simplex, and the affine transformation of a simplex is also a simplex, thus new endmembers can be determined sequentially [122].

After extracting MT-EMs in each patch, an *endmember pool* $\boldsymbol{U}$ is built, which is the union of all the MT-EM sets extracted from Z patches, i.e., $\boldsymbol{U} = \boldsymbol{E}_{S,1} \cup \boldsymbol{E}_{S,2} \cup \ldots \cup \boldsymbol{E}_{S,Z}$. Let P_s be the total number of MT-EMs in $\boldsymbol{U}$, $P_s = \sum_{z=1}^{Z} P_{s,z}$. Note that $\boldsymbol{U}$ represents all the distinct MT-EMs extracted at a local scale defined by patches. All P_s endmembers in $\boldsymbol{U}$ are used in the unmixing model at the global scale. The linear mixture model (LMM) is considered, which assumes a pixel is the result of a linear combination of endmember signatures weighted with their abundances [110]. LMM has been intensively investigated in the literature [109], [110], [111], [117]. However, as mentioned in Section II, the original LMM is built based on the single-date images, whereas here a multitemporal (stacked) domain is considered. Thus an approximation is made for a given pixel $x_s(i,j)$ ($1 \le i \le I$, $1 \le j \le J$) in $\boldsymbol{X}_S$ that its spectrum $\boldsymbol{x}_s(i,j)$ follows a linear mixture of a unique changed or unchanged MT-EMs, which is modeled as:

$$\boldsymbol{x}_s(i,j) = \sum_{p=1}^{P_s} a_p(i,j)\boldsymbol{e}_p(i,j) + \boldsymbol{n}(i,j) \tag{26}$$

where $\boldsymbol{e}_p$ is the spectral signature of the p-th (p=1,…, P_s) endmember in $\boldsymbol{X}_S$, a_p denotes the corresponding fractional abundance (which is the percentage of $\boldsymbol{e}_p$ within the considered pixel), and $\boldsymbol{n}$ is the noise vector. $\boldsymbol{A}_S$ and $\boldsymbol{U}$ are the set of a_p and $\boldsymbol{e}_p$ (p=1,…, P_s) for all pixels in $\boldsymbol{X}_S$, respectively.

Based on the extracted endmember pool $\boldsymbol{U}$, unmixing is conducted to estimate the abundances $\boldsymbol{A}_S$ of all MT-EMs in $\boldsymbol{U}$ by solving the following non-negative constrained least squares problem:

$$\begin{gathered} \tilde{\boldsymbol{A}}_S = \arg\min_{\boldsymbol{A}_S} \left\| \boldsymbol{X}_S - \boldsymbol{U}\boldsymbol{A}_S \right\|^2 \\ \text{subject to: } \boldsymbol{A}_S \ge 0 \end{gathered} \tag{27}$$

where $\boldsymbol{A}_S \ge 0$ is the imposed abundance nonnegative constraint (ANC).

6.2.3. Change Analysis

$\boldsymbol{U}$ includes all local MT-EMs that either belong to the set of change classes Ω_c or to the set of no-change classes Ω_n. Thus $\boldsymbol{U}$ can be divided into two subsets $\boldsymbol{U} = \{\boldsymbol{U}_c, \boldsymbol{U}_n\}$, where $\boldsymbol{U}_c$ and $\boldsymbol{U}_n$ indicates the endmember pool for Ω_c and Ω_n, respectively. Note that in this work we are only interested in distinguishing and identifying the K unique change classes $\left\{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}\right\}$ in $\boldsymbol{U}_c$, whereas we consider the Ω_n in $\boldsymbol{U}_n$ as one general no-change class ω_n, $\Omega_n \approx \omega_n$. Let $P_{s,c}$ and $P_{s,n}$ be the number of endmembers in $\boldsymbol{U}_c$ and $\boldsymbol{U}_n$, respectively, and P_s=$P_{s,c}$+$P_{s,n}$. The change analysis aims to separate two classes of MT-EMs (i.e., $\boldsymbol{U}_c$

and $\boldsymbol{U}_n$) and to identify the unique change classes in $\boldsymbol{U}_c$. To this end, the magnitude of the spectral difference (see Fig.3-5, two components in the extracted MT-EMs associated to $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively) is analyzed according to change vector analysis:

$$\rho_{e_p} = \sqrt{\sum_{b=1}^{B} \left(e_{p,B+b} - e_{p,b} \right)^2} \tag{28}$$

where $e_{p,b}$ is the b-th stacked channel of the p-th (p = 1,…, P_s) endmember in $\boldsymbol{U}$. A given $\boldsymbol{e}_p$ is classified either into $\boldsymbol{U}_c$ or $\boldsymbol{U}_n$ depending on its magnitude value, which can be higher or lower than T_ρ. The threshold T_ρ can be estimated manually or automatically based on the histogram of the magnitude variable ρ calculated on $\boldsymbol{X}_D$ as [92], [31], [86], [88]:

$$\rho = \sqrt{\sum_{b=1}^{B} \left(\boldsymbol{X}_{D,b} \right)^2} \tag{29}$$

where $\boldsymbol{X}_{D,b}$ is the b-th (b=1,...,B) component of $\boldsymbol{X}_D$. This can be done by using one of the techniques proposed in the literature [31], [88].

Since the MT-EMs in $\boldsymbol{U}_c$ are extracted from different patches representing the spatial distribution of different change targets on the images, a grouping step is required to cluster the MT-EMs that represent the same change class while showing its spectral variability. The final MT-EM groups contain the unique change classes represented at a global level. For the endmember grouping, we adopted an iterative scheme that was designed based on the Spectral Angle Mapper (SAM) [114]. Instead of using SAM, in this work we selected the spectral measurement proposed in [123], which considers both the spectral shape information by SAM and the stochastic behavior of the spectra by spectral information divergence (SID). Let $\boldsymbol{e}_\alpha$ and $\boldsymbol{e}_\beta$ be two given MT-EMs in $\boldsymbol{U}_c$. Let $\boldsymbol{r} = (r_1, r_2, \ldots, r_{2B})^T$ be the probability vector of $\boldsymbol{e}_\alpha = (e_{\alpha,1}, e_{\alpha,2}, \ldots, e_{\alpha,2B})^T$ with $r_b = e_{\alpha,b} \big/ \sum_{b=1}^{2B} e_{\alpha,b}$, and $\boldsymbol{m} = (m_1, m_2, \ldots, m_{2B})^T$ for $\boldsymbol{e}_\beta$ with $m_b = e_{\beta,b} \big/ \sum_{b=1}^{2B} e_{\beta,b}$. The SID measure is defined as [95]:

$$\mathrm{SID}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) = \sum_{b=1}^{2B} r_b \log(r_b / m_b) + \sum_{b=1}^{2B} m_b \log(m_b / r_b) \tag{30}$$

and the SAM is defined as:

$$\mathrm{SAM}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) = \cos^{-1}\left[\sum_{b=1}^{2B} e_{\alpha,b} e_{\beta,b} \Bigg/ \sqrt{\sum_{b=1}^{2B} (e_{\alpha,b})^2 \sum_{b=1}^{2B} (e_{\beta,b})^2} \right] \tag{31}$$

So the SID-SAM combined spectral measure ϑ is defined as [123]:

$$\theta(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) = \mathrm{SID}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) \cdot \sin\left[\mathrm{SAM}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) \right] \tag{32}$$

where sin(·) is the trigonometric sine function.

The MT-EM grouping starts from a random initialization, where the first class ω_{c_1} is randomly assigned to an endmember in $\boldsymbol{U}_c$. The SID-SAM measure θ is then computed between it and each of the remaining endmembers in $\boldsymbol{U}_c$. If the value of θ is smaller than a given threshold T_θ, the considered endmember is clustered into ω_{c_1}. Then the grouping procedure continues for the second class ω_{c_2} on those endmembers without a label. The iteration terminates when all endmembers have a label. Note that the defined threshold T_θ controls the similarity between endmembers required for the grouping. This is the only user-defined parameter in the proposed technique. Finally, the endmember grouping results in K unique kinds of changed MT-EMs $\left\{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}\right\}$. By considering also the no-change class ω_n, in total we have $K'=K+1$ classes of MT-EMs in $\boldsymbol{X}_S$.

6.2.4. Abundance Combination and Final Change-Detection Map Generation

Based on the endmembers grouping result, the abundances of local MT-EMs that belong to a given class are summed together, thus generating the final abundance for that class. Let ω_ε be a given class in $\Omega=\left\{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}, \omega_n\right\}$. The final abundance map $\boldsymbol{A}_{S,\varepsilon}$ of the class ω_ε for all pixels in $\boldsymbol{X}_S$ is computed as:

$$\boldsymbol{A}_{S,\omega_\varepsilon} = \sum_{\boldsymbol{e}_p \in \omega_\varepsilon} \boldsymbol{A}_{S,p} \tag{33}$$

where $\boldsymbol{A}_{S,p}$ is the abundance map of a given MT-EM $\boldsymbol{e}_p$ in $\boldsymbol{U}$, and $\boldsymbol{e}_p \in \omega_\varepsilon$, $p=1,\ldots, P_s$.

Thus the considered multiple-change detection problem can be solved by assigning the final class label of a pixel $x_s(i, j)$ ($1 \le i \le I$, $1 \le j \le J$) in $\boldsymbol{X}_S$ to the class ω_ε having the maximum abundance value:

$$x_s(i,j) \in \underset{\omega_\varepsilon \in \Omega}{\arg\max}\left(a_{\omega_\varepsilon}(i,j)\right) \tag{34}$$

where $a_{\omega_\varepsilon}(i,j)$ is the abundance value of class ω_ε in pixel $x_s(i,j)$. Note that anyway the derived abundance maps can be considered as sub-pixel change-detection results.

6.3 Data Set Description and Design of Experiments

Data Set 1: Simulated hyperspectral remote sensing data set

The first data set is made up of a hyperspectral image acquired by the AVIRIS sensor in 1998 on Salinas valley, California. The original image has 224 contiguous spectral bands with wavelength from 400nm to 2500nm, characterized by a spatial resolution of 3.7m and a spectral resolution of 10nm. Ground truth data are available that contain 16 material classes (e.g., vegetation, bare soil, and vineyard), thus usually this data set is used for hyperspectral image classification. A subset was selected on the whole image having a size of 217×97 pixels. In the pre-processing, 20 water absorption bands (i.e., bands 108-112, 154-167 and 224) were discarded, obtaining 204 bands for our experiments. Taking advantage of the available ground truth, we simulated an image (considered as $\boldsymbol{X}_2$) based on the original image (considered as $\boldsymbol{X}_1$). In order to obtain a realistic simulated image, ten tiles were extracted from $\boldsymbol{X}_1$ (see Fig.6-2.a) and inserted back in different areas on $\boldsymbol{X}_1$ by replacing the whole spectral vectors. Thus $\boldsymbol{X}_2$ was generated with eight simulated change classes (see Fig.6-2.b). A small constant bias value was applied to $\boldsymbol{X}_2$ to simulate a stationary radiometric change. White Gaussian noise was added to $\boldsymbol{X}_2$ with different levels of Signal-to-noise ratio (SNR) values (i.e., SNR=10, 20, 30, 40 dB). Thus, we obtained four image pairs built by $\boldsymbol{X}_1$ and one out of the four simulated $\boldsymbol{X}_2$. False color composites of $\boldsymbol{X}_1$ and one of the simulated $\boldsymbol{X}_2$ (i.e., SNR=20) are shown in Fig.6-2 (a) and (b), respectively. Fig.6-2 (c) is the reference change map. Detailed simulated land-cover transitions are listed in TABLE 4-1, which provides the number of samples for each simulated change class. Note that we did not simulate mixture at single pixel level but we used this data set for assessing the effectiveness of the proposed method when a large number of changes are present in noisy images.

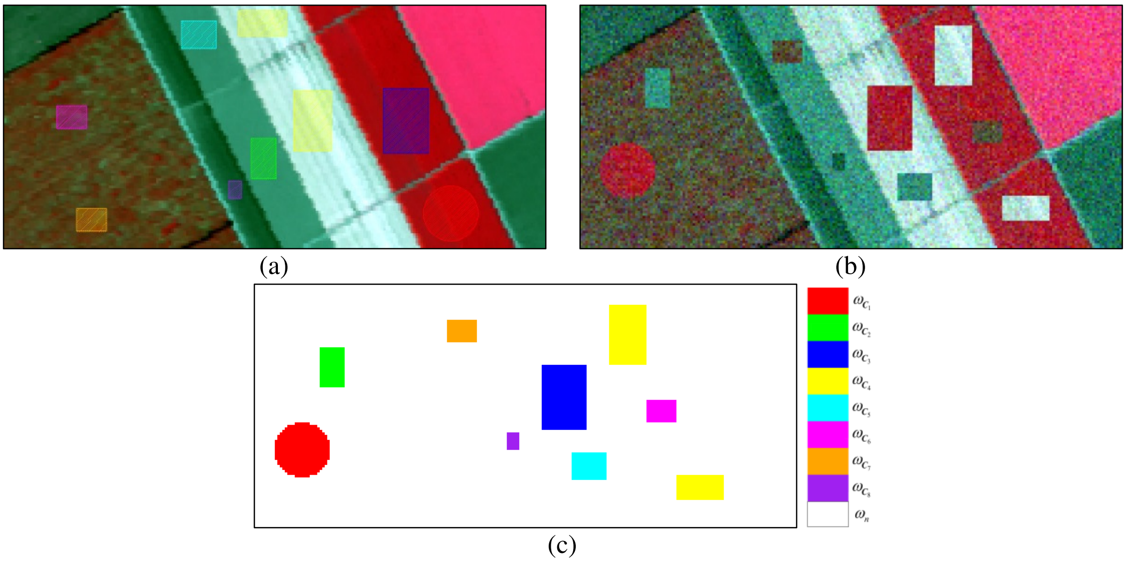


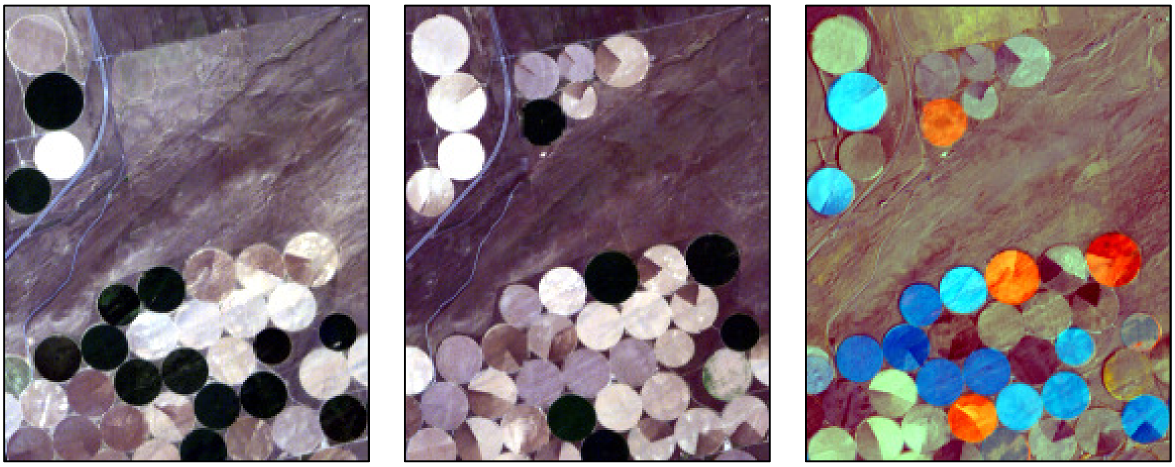
Fig.6-2 False color composites (Bands: R: 40, G: 30, B: 20) of (a) the hyperspectral image acquired by the AVIRIS sensor in Salinas scenario ($\boldsymbol{X}_1$) and (b) the simulated changed image ($\boldsymbol{X}_2$) computed with an additive white Gaussian noise (SNR=20 dB). (c) Change reference map (eight changes in different colors, and no-change class in white color).

TABLE 6-1 SIMULATED CHANGE CLASSES AND RELATED NUMBER OF SAMPLES (SALINAS DATA SET)

Change class	**Simulated changes (from $\boldsymbol{X}_1$ to $\boldsymbol{X}_2$)**	**Samples (Number of pixels)**
ω_{C1}	Celery → Vinyard_untrained	388
ω_{C2}	Fallow_smooth → Vinyard_untrained	160
ω_{C3}	Celery → Stubble	468
ω_{C4}	Stubble → Celery	550
ω_{C5}	Fallow_smooth → Stubble	154
ω_{C6}	Vinyard_untrained → Celery	108
ω_{C7}	Vinyard_untrained → Fallow_smooth	108
ω_{C8}	Fallow_rough_plow → Fallow_smooth	35
ω_n	No-change	19078

Data Set 2: Real Hyperion hyperspectral remote sensing data set

This data set is made up of a pair of real bi-temporal hyperspectral remote sensing images acquired by the Hyperion sensor mounted onboard the EO-1 satellite on May 1, 2004 ($\boldsymbol{X}_1$) and May 8, 2007 ($\boldsymbol{X}_2$). The study area is an agricultural irrigated land of Umatilla County, Oregon, United States, which has a size of 180×225 pixels. The Hyperion image has a wavelength range from 350nm to 2580nm, characterized by a spectral resolution of 10nm and a spatial resolution of 30m. After the pre-processing phase (i.e., bad stripes repairing, uncalibrated and noisiest bands removal, atmospheric correction, co-registration), 159 bands (i.e., bands: 8-57, 82-119, 131-164, 182-184, 187-220) out of the original 242 bands were used for the considered CD task. Occurred changes in this scenario mainly include the land-cover transitions between crops, bare soil, variations in soil moisture and water content of vegetation. Note that no ground truth data are available for this data set, thus the detailed validation of the results was done qualitatively through a careful visual analysis. Fig.6-3 (a) and (b) show the false color composite of the $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ images, respectively. Fig.6-3 (c) presents a false color composite of three bands in $\boldsymbol{X}_D$, thus possible change classes are shown in different colors whereas the gray pixels indicate the no-change background.



(a) (b) (c)

Fig.6-3 Bi-temporal Hyperion images acquired on an agricultural irrigated scenario. False color composite (wavelength: R: 650.67nm, G: 548.92nm, B: 447.17nm) images acquired in: (a) 2004 ($\boldsymbol{X}_1$), (b) 2007 ($\boldsymbol{X}_2$), and (c) composite of three $\boldsymbol{X}_D$ channels (wavelength: R: 823.65nm, G: 721.90nm, B: 620.15nm).

6.4 Experimental Results

6.4.1 Simulated Hyperspectral Data Set

Let us analyze in detail the experimental results obtained by the proposed MSU approach applied to one simulated image pair (associated with SNR=20dB). First, we generated $\boldsymbol{X}_S$, and divided it into four regular patches $\boldsymbol{X}_{S,1}$-$\boldsymbol{X}_{S,4}$ (Z=4), see Fig.6-4. In each patch, the number of MT-EMs was estimated by using the ELM algorithm. The VCA was implemented to extract MT-EMs. Thus $\boldsymbol{U}$ was obtained as the union of all MT-EMs from all patches (see Fig.6-4). The abundances $\boldsymbol{A}_S$ were then calculated according to (27). T_ρ was automatically estimated on the magnitude variable ρ [31] and was set equal to 0.82 to separate $\boldsymbol{U}_c$ and $\boldsymbol{U}_n$. T_ϑ (which is the only manually defined parameter) was set equal to 0.01 for controlling endmember grouping in $\boldsymbol{U}_c$. MT-EM abundances were summed into the final abundances of change and no-change classes based on the grouping results to generate the final change detection map (33), (34).

From Fig.6-4 we can see that the MSU approach extracted all simulated MT-EMs (i.e., P_s=35) from the four divided patches, representing very well the distinct endmembers (including both the change and no-change classes) in the considered local patches. After change analysis, all the simulated change targets were identified correctly having at least one hit MT-EM on each of them. Even the small size change classes were detected successfully. The endmember grouping resulted in eight unique change classes, which is the correct number of the simulated changes. The spectral signatures of the detected MT-EMs (associated with the detected eight change classes and the no-change class) are shown in Fig.6-5, and the corresponding abundances of each class are illustrated in Fig.6-6 (a)-(j).

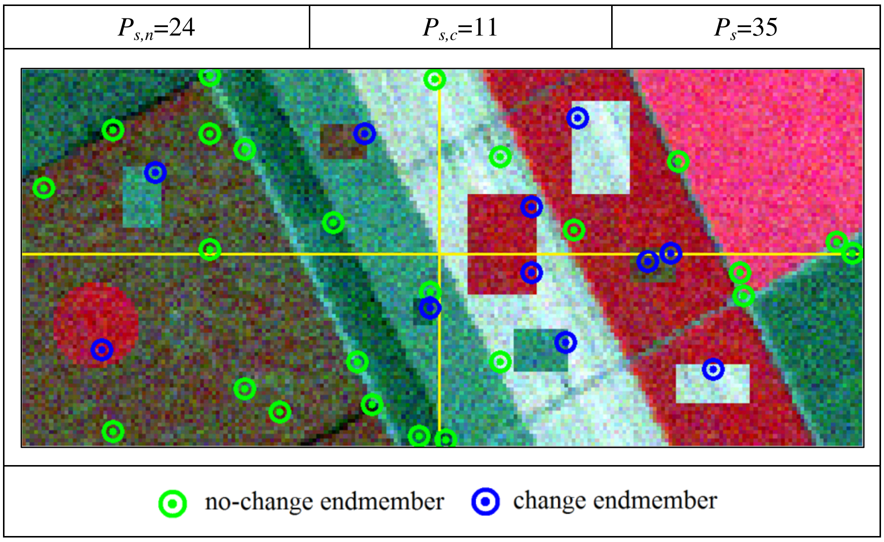


Fig.6-4 Extracted MT-EMs by the proposed MSU approach on the simulated data set with SNR=20dB. From left to right, up to down the four patches $X_{S,1}$ ($P_{s,1}$=10), $X_{S,2}$ ($P_{s,2}$=8), $X_{S,3}$ ($P_{s,3}$=9) and $X_{S,4}$ ($P_{s,4}$=8) are shown.

From Fig.6-5 and Fig.6-6, we can observe that: 1) The detected eight change classes have unique spectral signatures in the $\boldsymbol{X}_S$ domain that are discriminable among each other. The two components of MT-EM spectra have different shapes, indicating the change nature of the endmembers, whereas the no-change class has similar spectral shapes in two components; 2) MT-EMs identified from spatial different regions in $\boldsymbol{X}_S$ but grouped into a same class (e.g., see MT-EMs in change class 3, class 4 and class 6) show slight differences in their spectra thus confirming the endmember spectral variability. Their abundances are summed into the corresponding grouped class rather than being used to represent a class independently, thus to better describe the related unique change class present in the global scene; 3) the abundance maps show a good unmixing and separation result among different classes (see Fig.6-6.a-i), where the detected eight change targets and the no-change background shows a clear contrast of their abundances (with respect to the represented colors). Thus the considered multiple change detection problem was successfully solved by estimating the percentage of class substances in the subpixel level.

CD maps obtained by the proposed MSU approach on the four simulated hyperspectral data sets are shown in Fig.6-7. From a qualitative analysis of the CD maps, we can observe that the proposed MSU technique achieved good detection results under different noise levels. All eight change classes were successfully detected in all cases but the one with the highest noise level, where the small ω_{C8} was not detected (see Fig.6-7, SNR=10dB). This is due to the fact that as expected the significance of small changes decreases when increasing the noise levels, thus a 4-patch division is not sufficient to detect all of them. In this case, a smaller patch scale is required to reach the optimal detection scale.

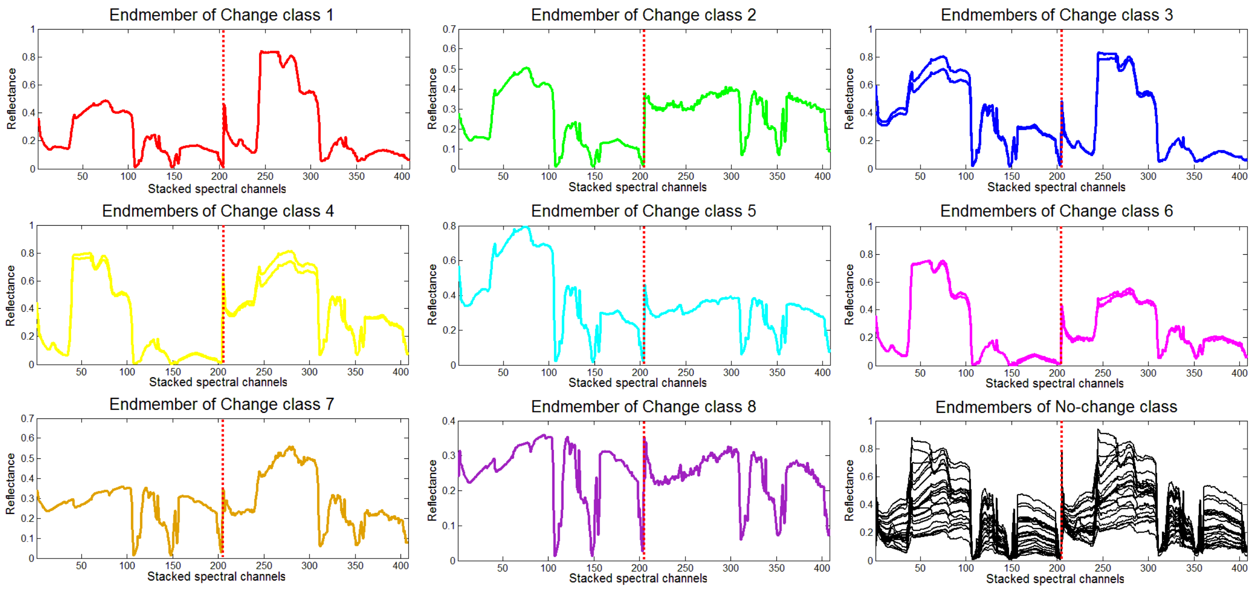


Fig.6-5 Spectral signatures of the MT-EMs extracted by the proposed MSU approach on the simulated data set (SNR=20 dB), where eight unique change classes (in different colors) are identified, and the no-change class endmembers are in black. Endmembers that belong to the same class represent the variability in the image.

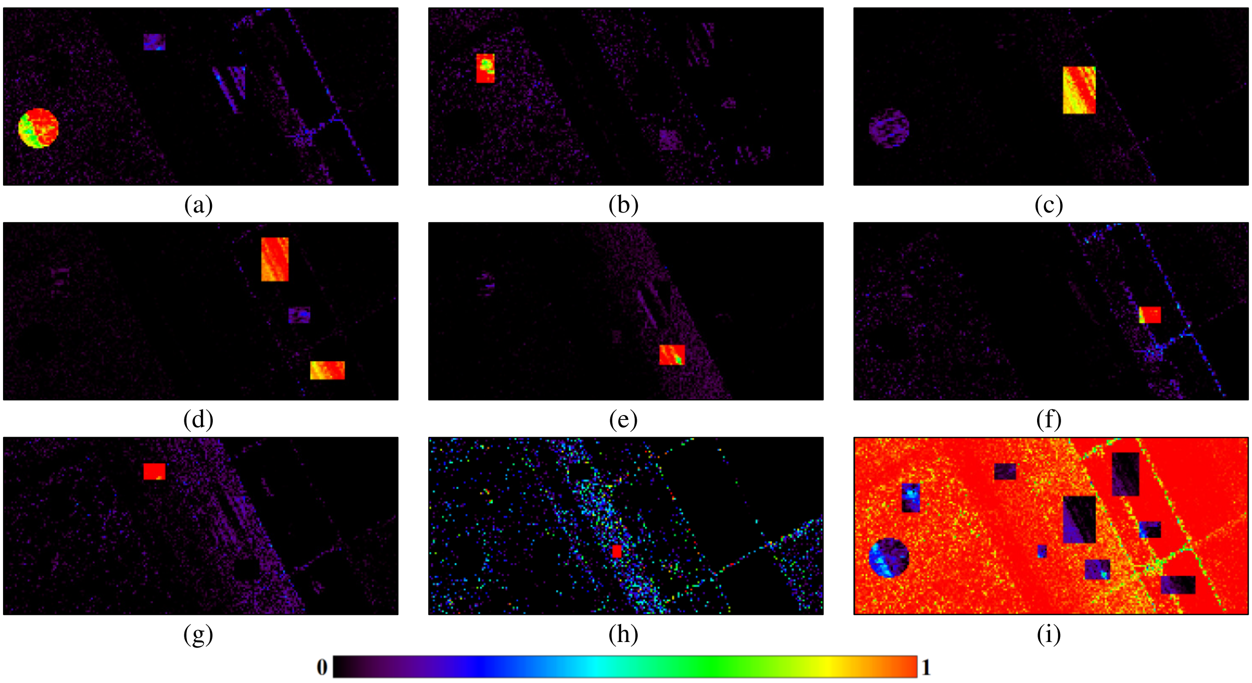


Fig.6-6 Final class abundances on the simulated data set (SNR=20dB), where (a)-(h) are the abundances of eight change classes, and (i) of the no-change class.

A detailed quantitative analysis was carried out by comparing the experimental results obtained by the proposed subpixel-level MSU approach with those yielded by three pixel-level CD-HS methods: 1) the hierarchical spectral change vector analysis (HSCVA) [86], 2) the sequential spectral change vector analysis (S^2CVA) [87], and 3) the unsupervised k-means clustering applied to the changed spectral change vectors. Experiments were carried out on the four simulated image pairs with different SNR values (i.e., 10-40 dB). Note that in the case of SNR=10dB the 8-patch (Z=8) scheme was used in the MSU approach, thus obtaining the detection of all eight change classes. Advantage was given to the k-means by giving as input the known number of classes (i.e., K=8), whereas for the other methods this number was estimated in the CD process. The final result of k-means was generated by the average of 100 trails in order to reduce the uncertainty of random initialization.

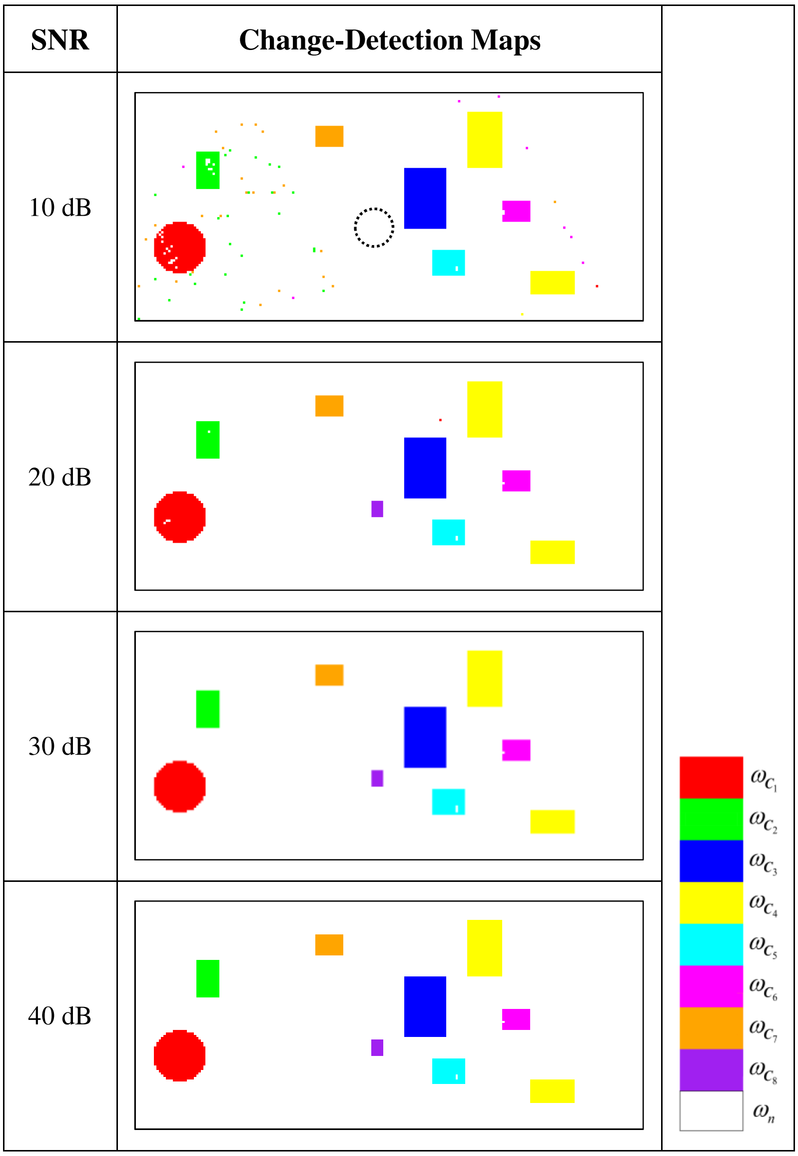


Fig.6-7 Change-detection maps obtained by the MSU approach on different simulated hyperspectral data sets (with SNR=10, 20, 30, 40 dB). Different change classes are in different colors, and the no-change class is in white color.

The numerical results are given in TABLE 6-2, where accuracy indices including the Overall Accuracy (*OA*), the Kappa Coefficient (*Kappa*) and the number of detection errors were computed according to the available reference map. Note that all the considered methods detected the correct number of changes (i.e., K=8). From TABLE 6-2, one can see that the two state-of-the-art hierarchical methods achieved good results in all four cases, which obtained the highest *OA* and *Kappa* values. The systematic top-down structure in the hierarchical analysis gradually recovers and models the hidden change information in the data set, thus resulting in more accurate results when compared with the proposed one-step processing [86], [87]. However, it is worth noting that: 1) both HSCVA and S^2CVA methods are designed in a semi-automatic fashion (i.e., an initialization for model selection in HSCVA [86] and a user interaction for change identification in S^2CVA [87] are required, respectively); 2) Effort is required to search for the hierarchical structure, which increases the implementation complexity and the time cost. For example, for the simulated data set with SNR=20dB, to successfully detect all eight change classes, both methods resulted in a three-level hierarchy with more than twelve nodes. On the contrary, the proposed automatic MSU method obtained only slightly worst CD result compared with the two state-of-the-art semi-automatic techniques without exploiting a hierarchical structure. Note that although the reference k-means algorithm was applied using the known number of classes, it resulted in a significantly higher number of errors than the other three methods (see TABLE 6-2). This indicates the difficulty of the CD-HS problem to be solved when applying the clustering methods directly to the high-dimensional hyperspectral images.

TABLE 6-2 CHANGE DETECTION ACCURACY AND ERROR INDICES OBTAINED BY THE CONSIDERED METHODS (DIFFERENT SIMULATED HYPERSPECTRAL DATA SETS).

SNR	Method	Unsupervised and Automatic Detection?	*OA* (%)	*Kappa*	Errors (pixel)
10 dB	HSCVA	No	99.69	0.9829	64
	S^2CVA	No	99.74	0.9856	54
	k-means	No	98.32	0.8988	353
	MSU (Z=8)	Yes	99.54	0.9749	96
20 dB	HSCVA	No	99.99	0.9995	2
	S^2CVA	No	99.98	0.9989	4
	k-means	No	99.47	0.9702	111
	MSU (Z=4)	Yes	99.96	0.9978	8
30 dB	HSCVA	No	99.99	0.9997	1
	S^2CVA	No	99.98	0.9989	4
	k-means	No	99.66	0.9811	71
	MSU (Z=4)	Yes	99.98	0.9989	4
40 dB	HSCVA	No	99.99	0.9997	1
	S^2CVA	No	99.98	0.9989	4
	k-means	No	99.78	0.9880	45
	MSU (Z=4)	Yes	99.99	0.9992	3

6.4.2 Real Hyperion Remote Sensing Satellite Data Set

Experiments were carried out by using the proposed MSU approach and the three pixel-level reference methods (i.e., HSCVA, S^2CVA, *k*-means) as in the previous case. $\boldsymbol{X}_S$ was divided into four regular patches (Z=4) (see Fig.6-8). The ELM algorithm was applied to each local patch to identify the number of MT-EMs (i.e., Fig.6-8 $P_{s,1}$-$P_{s,4}$). A total of 42 MT-EMs were identified, which consider the local spectral distinct endmembers thus better describing the spectral composition in the whole scene. The VCA was used to extract local MT-EMs in each patch. All the extracted MT-EMs are shown in Fig.6-8. The abundances $\boldsymbol{A}_S$ were calculated according to (26) and (27). The change analysis was done by automatically estimating the T_ρ equal to 1.486 to separate $\boldsymbol{U}$ into $\boldsymbol{U}_c$ and $\boldsymbol{U}_n$. The only parameter to be manually selected for the MT-EM grouping in $\boldsymbol{U}_c$ was defined as T_ϑ =0.01. Seven unique change classes were thus detected. The final class abundances were then summed and compared in each pixel of $\boldsymbol{X}_S$ to define the final change detection map according to (33) and (34).

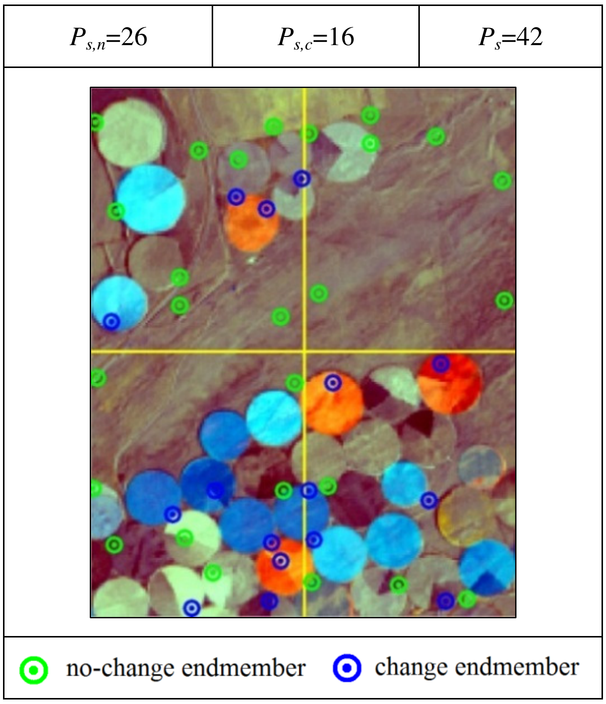


Fig.6-8 MT-EMs extracted by the proposed MSU approach on the real bitemporal Hyperion data set. From left to right, up to down the four split patches of $X_{S,1}$ ($P_{s,1}$=12), $X_{S,2}$ ($P_{s,2}$=7), $X_{S,3}$ ($P_{s,3}$=13) and $X_{S,4}$ ($P_{s,4}$=10) are shown.

The spectral signatures of the extracted MT-EMs by MSU are illustrated in Fig.6-9, including seven change classes in different colors and the no-change class in black. Their corresponding abundances are shown in Fig.6-10 (a)-(h). We can observe that: 1) different or similar spectral shapes in two components of the MT-EM spectrum indicate the presence of the change classes and the no-change class in the $\boldsymbol{X}_S$ domain, respectively. The MT-EMs of the detected change classes have discriminable and unique spectral signatures among each other (See Fig.6-9 change classes 1-7); 2) an endmember set rather than a single endmember was used to represent a detected change class, thus the occurred change targets were described via unmixing by also considering the endmember variability (see Fig.6-9 change classes 1-3, 5-7); 3) we are not interested in distinguishing MT-EMs of different no-change classes among each other in $\boldsymbol{U}_n$. However, as an important and non-negligible source of endmembers in the defined mixture model in $\boldsymbol{X}_S$, they were considered in the unmixing process, and their abundances were summed into the final no-change class ω_n. A good estimation of the no-change background can be seen in Fig.6-10 (h), which is well separated from the other change classes; 4) the abundances of change classes (see Fig.6-10 a-g) confirm the accurate representations of the unique changes and a discrimination among them. From the abundance maps, one can easily observe the spatial distribution of different change classes in the scene and investigate in detail their composition within a pixel, thus better understanding and solving the considered CD problem at subpixel level. Note that the circle patterns identified in the change class 7 in Fig.6-10 (g) might be associated to the misregistration errors of roads that surround the agricultural irrigated fields. Visual analysis on the Google maps [124] tells us that those roads have an average width between 6m to 10m (see the screenshots of google maps shown in Fig.6-11), which is less than half pixel on the Hyperion images whose spatial resolution is 30m. Despite in the pre-processing step the residual value of co-registration is limited within 0.5 pixel, the co-registration errors may still contribute to the definition of change endmembers and thus change classes. However, these classes can be associated with noise in a post-processing analysis, and possibly used for the optimization of the co-registration process [125].

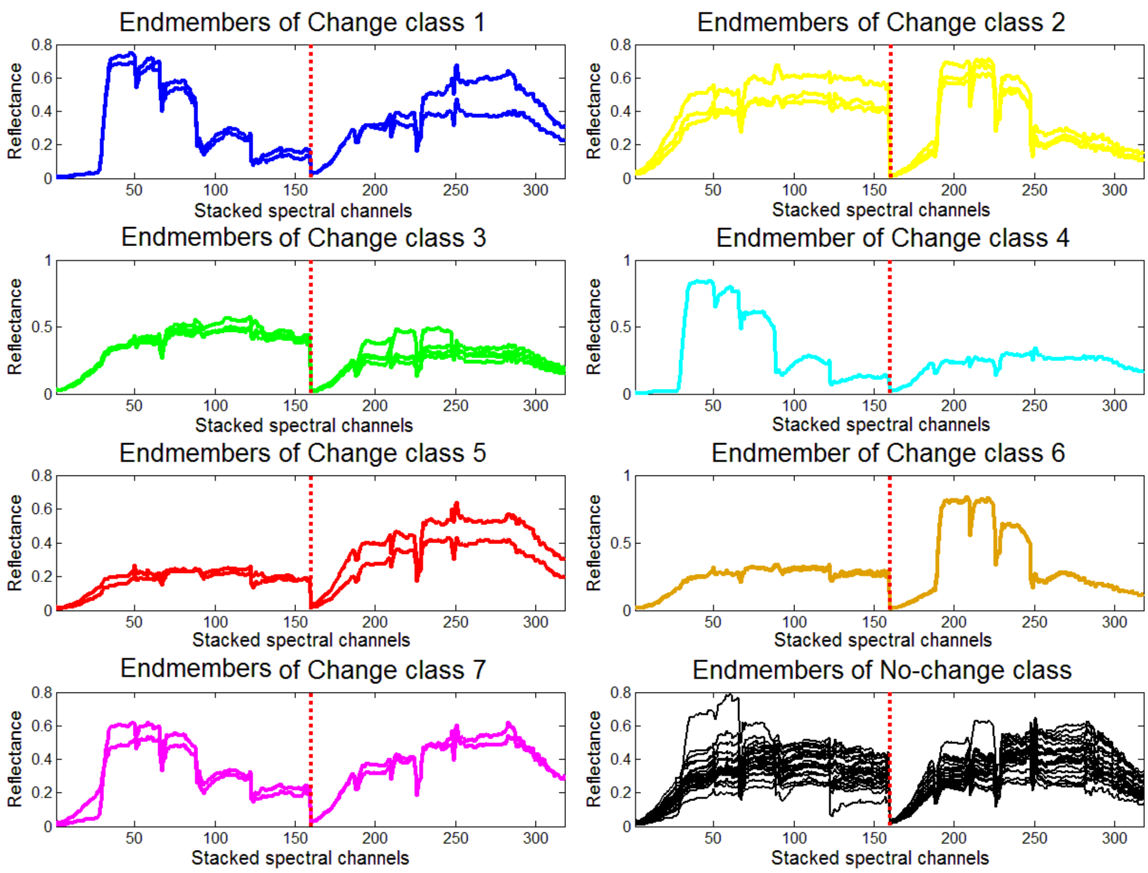


Fig.6-9 Spectral signatures of MT-EMs extracted by the proposed MSU approach on the real bi-temporal Hyperion remote sensing data set. Seven unique change classes are identified and shown in different colors, while the no-change MT-EMs are in black. Endmembers that belong to the same class represent the variability in the image.

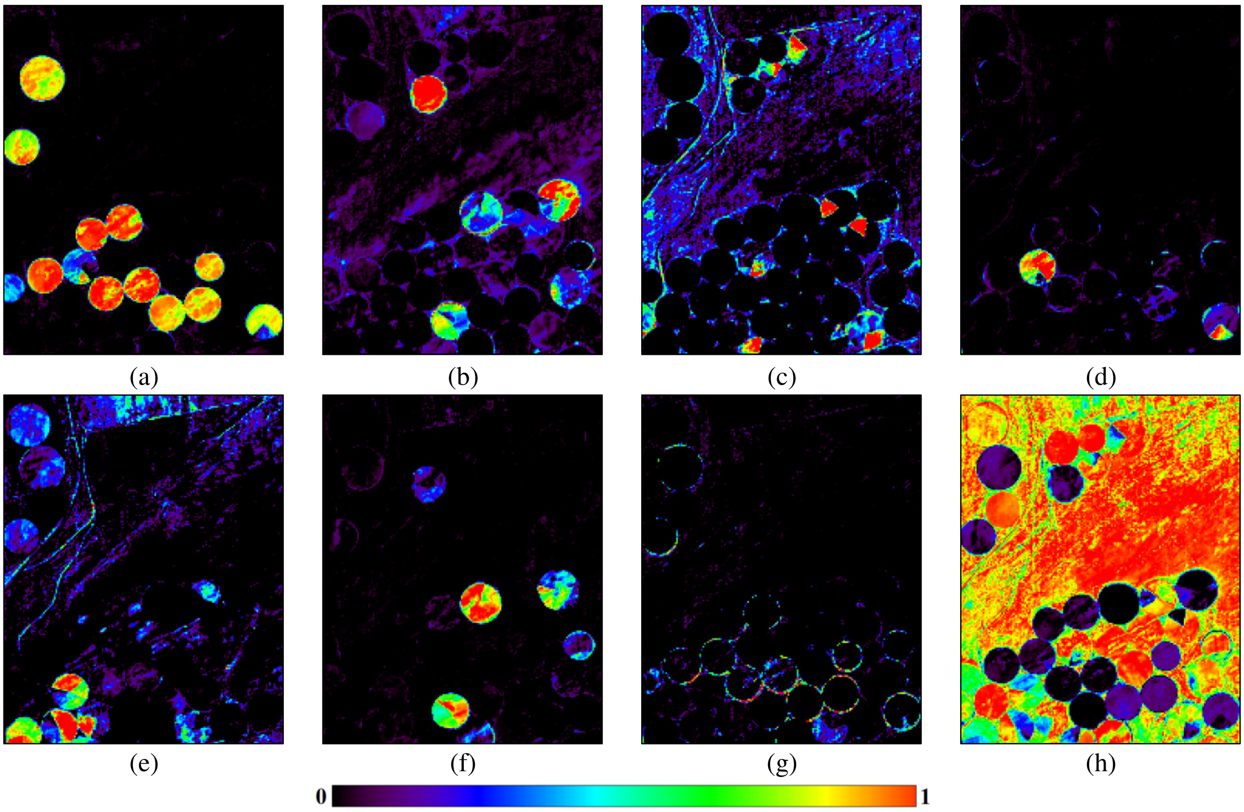


Fig.6-10 Final class abundances obtained by MSU on the real bi-temporal Hyperion data set: (a)-(g) show the seven change classes and (h) the no-change class.

Fig.6-11 Circle roads that surround the agricultural irrigated fields observed from Google Maps [124]

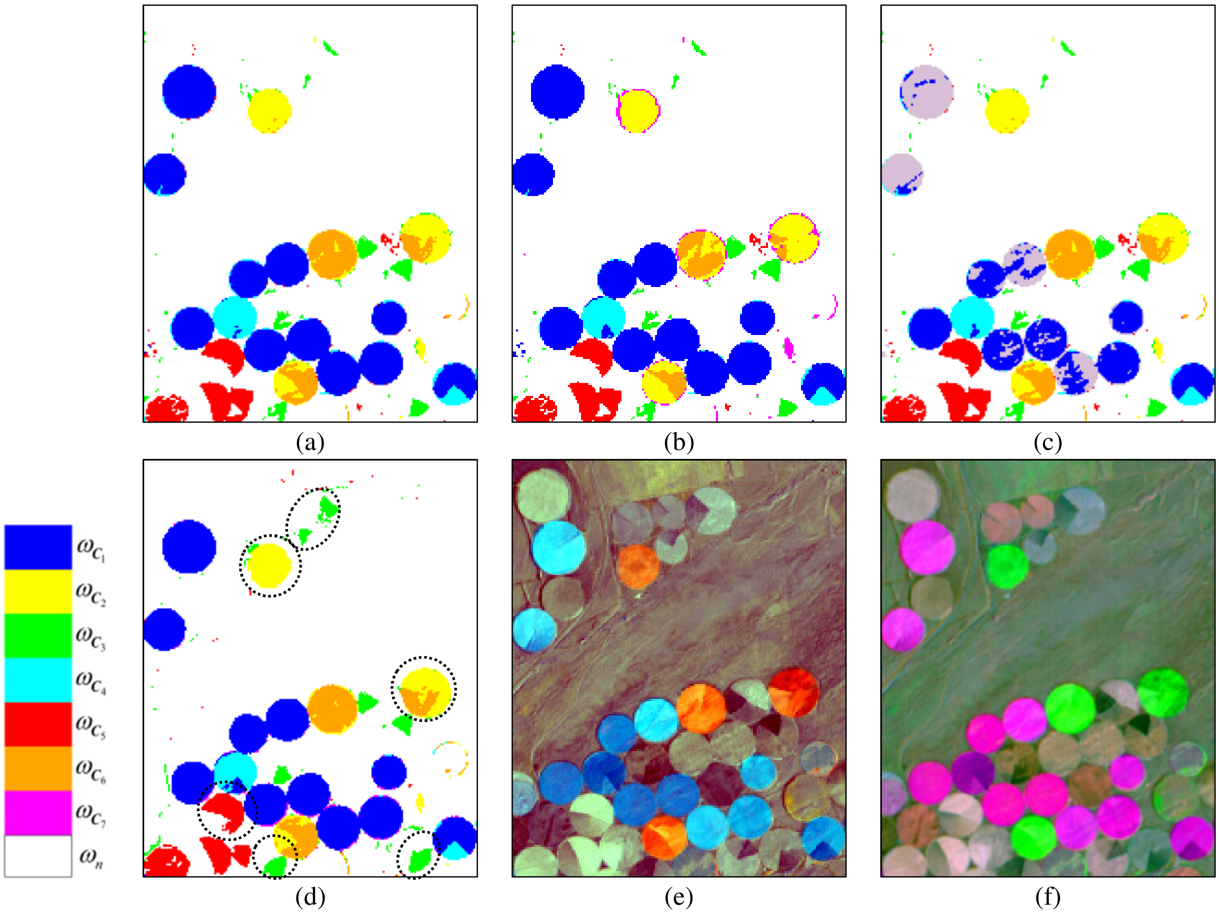


Fig.6-12 Change-detection maps obtained by: (a) HSCVA [86] (K=6); (b) S2CVA [87] (K=7); (c) k-means (K=7) and (f) the proposed MSU (K=7) on the real bitemporal Hyperion data set. Different changes are in different colors and the no-change class is in white color. Two reference false color composites of XD are provided for visual comparison in: (d) wavelength: R: 823.65nm, G: 721.90nm, B: 620.15nm; (e) R: 1729.7nm, G: 752.43nm, B: 548.92nm.

Change-detection maps obtained by the considered approaches are shown in Fig.6-12 (a)-(d). The two reference hierarchical methods detected six and seven kinds of changes, respectively. The proposed MSU detected K=7 change classes as S^{2}CVA. It can be clearly seen that change targets were better recognized in the result of MSU (highlighted in Fig.6-12.d) than in those of the two hierarchical methods (see Fig.6-12 a and b) and the *k*-means (see Fig.6-12.c). Comparing to the reference false color composite of $\boldsymbol{X}_D$ bands in Fig.6-12 (e) and (f), one can observe that the changes highlighted in Fig.6-12.d (especially the change class ω_{c_3} in green color) were detected more accurately. Accordingly, the multitemporal spectral mixture model used by the proposed MSU approach investigating in details the spectral composition of a pixel in $\boldsymbol{X}_S$, avoids the propagation of errors due to the crisp modeling of change analysis accomplished by the reference methods. Despite an advantage was given to the *k*-means method providing as input the same class number obtained by MSU (i.e., K=7), more fragments are present in the detected change map (see Fig.6-12.c). It is important to note that the proposed MSU approach was implemented in an automatic and unsupervised way. However, it detected successfully the multiple change classes and resulted in comparable (slightly better) CD results than the other techniques without any manual operation as those required in the two semi-automatic reference methods.

6.5 Discussion and Conclusions

In this chapter, a novel multitemporal spectral unmixing (MSU) approach has been proposed to address the challenging multiple change detection problem in multitemporal hyperspectral images. The proposed method is designed in an automatic and unsupervised way, thus is independent from the availability of prior knowledge and the manual assistance of the user in the real applications. The main novelties and contributions of the proposed method are as follows: 1) it provides a new perspective to detect changes by jointly exploring the spectral-temporal variations in $\boldsymbol{X}_S$ (i.e., spectral stacked domain); 2) it proposes a multitemporal spectral unmixing framework to solve the multiple change detection problem, where the identification of the number of change classes is done by identifying the distinct endmembers and the unique change classes, and the discrimination of changes is addressed by unmixing and abundances analysis; 3) it allows one to understand in details the spectral composition of a pixel, thus implementing CD at subpixel level. Experimental results obtained on both simulated and real hyperspectral images confirmed the effectiveness of the proposed MSU approach.

From the theoretical analysis and the practical experimental results, we can conclude that:

1) By taking advantage of the endmembers extraction and spectral unmixing while considering endmember variability (i.e., local endmembers strategy), the proposed method models well the change and no-change spectral compositions inside of a pixel. A more reliable decision is made according to the analysis of the endmember abundances associated with a given class with respect to a crisp decision based on the pure-pixel theory. Accordingly, more subpixel level spectral variations are expected to be identified, which are usually not detectable in the pixel-level-based state-of-the-art techniques.

2) The proposed method is automatic and unsupervised, thus allowing a fast and efficient CD without requiring any manual assistance of the users. Note that other endmember extraction techniques or unmixing models can be easily adopted within the proposed MSU approach.

Future developments of this work will include: 1) a joint analysis of the spatial-spectral-temporal information and the related impact on the proposed method; and 2) the extension of the proposed approach to the hyper-temporal (i.e., time series) domain.

Chapter 7

Conclusions

This chapter draws the conclusion of the research activities carried out in this thesis. It also summarizes and discusses the results. Finally, the future developments are described.

In this thesis, the interesting but challenging multiple-change detection problem in multitemporal hyperspectral images is comprehensively analyzed. Advanced techniques for automatic change detection in hyperspectral images have been developed. We: 1) intensively reviewed the literature on CD techniques in multitemporal multispectral and hyperspectral images; 2) analyzed the problems and challenges when dealing with CD on hyperspectral images; 3) proposed novel concepts for modelling the considered CD problem in the spectral difference domain and the multitemporal spectral stacked domain; 4) developed a semi-automatic technique for iteratively discovering, representing and detecting multiple changes; 5) proposed a partially-unsupervised hierarchical clustering approach to identify multiple changes; 6) proposed an automatic multitemporal spectral unmixing technique for multiple change detection.

A comprehensive overview of the state-of-the-art change-detection techniques is presented for both multitemporal multispectral and hyperspectral images. Limitations and challenges of the literature techniques have been analyzed and discussed. Then a qualitative and theoretical analysis of the relevant concepts in two change representation domains has been presented. It provides an empirical but solid analysis for the formulization of the CD-HS task, which leads to the proposal of corresponding solutions in the thesis.

In the *spectral difference domain*, by taking into account the intrinsic complexity of the hyperspectral data, the concepts of “major change” and “subtle change” have been defined from the perspective of pixel spectral behaviours, along with a deep analysis of their structures. This resulted in the definition of a hierarchical modelling of the considered complex CD-HS problem. Based on this analysis, we proposed two novel solutions.

The first solution is a novel sequential spectral change vector analysis (S^2CVA) method developed for discovering and detecting multiple changes (i.e., different kinds of changes) in multitemporal hyperspectral images. The proposed method is designed to be sensitive to the small spectral variations that can be identified in hyperspectral images but usually are not detectable in multispectral images. The proposed method exploits an iterative hierarchical scheme that at each iteration discovers and identifies a subset of changes. Developed on the state-of-the-art C^2VA representation, the proposed Adaptive Spectral Change Vector Representation (ASCVR) method adaptively and automatically changes the reference vectors according to the SCVs associated to the specific changes that are analyzed. Therefore, although the compression from the *B*-Dimensional to the 2-Dimensional feature space introduces an unavoidable loss of information, the sequential analysis gradually recovers in the hierarchy the information loss at first levels, resulting in a complete change representation and in a satisfactory change-detection map. The proposed approach is developed in an interactive and semi-automatic fashion, which allows one to study in detail the structure of changes hidden in the variations of the spectral signatures according to a top-down procedure. Despite the iterative nature, the computational complexity of the proposed method is very low. Few minutes were required for the entire processing on a standard PC for all the experimental data sets. In general, the total processing time depends on the size of the images and also on the depth of the hierarchical tree (the complexity of the changes present in the data). Experimental results obtained on three hyperspectral data sets confirmed the effectiveness of the proposed method.

The second solution is a novel hierarchical spectral change clustering approach, which aims to identify the possible changes occurred between a pair of hyperspectral images. Changes having discriminable spectral behaviors in hyperspectral images are identified hierarchically by considering coarse to fine spectral change significance in different levels. This leads to a better model of the complex structure of changes in the hyperspectral images, which is usually ignored by the state-of-the-art techniques based on a single level analysis. In particular, the proposed approach models the detection of multiple changes according to an iterative clustering, thus the separation of multiple change information (i.e., major changes, subtle changes and finally change endmembers) is carried out according to the difference of their SCVs represented in the spectral difference domain. In this way, we progressively decompose the complex problem into several specific sub-problems, and focus on each single portion of the multiple-change information highlighted in a given detection level. This makes it possible to effectively discover the difference among similar changes by decreasing the complexity of detection. Moreover, the proposed approach is designed in a partially-unsupervised fashion, thus it fits most of actual applications, for which often the ground truth data are not available. Satisfactory experimental results obtained on simulated and real bi-temporal images confirm the validity of the proposed approach with higher change-detection overall accuracy and higher multi-class separability measurement.

In the *multitemporal spectral stacked domain*, a multitemporal spectral mixture model has been defined to formalize the same CD problem from the point of view of spectral composition within a pixel. A novel multitemporal spectral unmixing (MSU) approach has been proposed for addressing the challenging multiple-change detection problem in bi-temporal hyperspectral images. Differently from the state-of-the-art methods that are mainly designed at pixel level, the proposed approach investigates in detail the spectral-spatial-temporal variations at subpixel level. The considered CD problem is analyzed in a multitemporal spectral stacked domain, where a bitemporal spectral mixture model is defined to formalize and analyze the spectral composition within a pixel. Different distinct multitemporal endmembers (MT-EMs) are extracted according to an automatic unmixing technique. Then a change analysis strategy is designed to distinguish the change and no-change MT-EMs. An endmember grouping scheme is applied to the changed MT-EMs to detect the unique change classes present in the scene. Finally, the considered multiple-change detection problem is solved by analyzing the abundances of the change and no-change classes and their contribution to each pixel. In this way, a more reliable decision is made according to the analysis of the endmember abundances associated to a given class rather than a crisp decision based on the pure-pixel theory. Thus more spectral variations are expected to be identified at subpixel level, which are usually not detectable in the pixel-based literature techniques. The proposed technique is designed in an automatic and unsupervised way, thus is independent from the prior knowledge and the manual assistance of the user in the real applications. The proposed approach has been validated on both simulated and real multitemporal hyperspectral datasets presenting multiple changes. Experimental results confirmed the effectiveness of the proposed method.

In order to continue but extend the research activities carried out in this thesis, some remaining open issues and interesting topics will be considered as future developments: 1) study of effective change variables constructed from the high-dimensionality feature space to enhance the performance of change representation and discovery in a low-dimensionality compressed feature space; 2) analysis on the joint use of the spatial-spectral-temporal information to increase the robustness and the accuracy of the proposed techniques; 3) design of advanced semi-automatic/fully-automatic techniques for real change detection applications; 4) extension of the proposed approaches to the analysis of time series hyperspectral images.

List of Publications

BOOK CHAPTER

[B1] L. Bruzzone, **S. Liu**, F. Bovolo, P. Du., Change Detection in Multitemporal Hyperspectral Images, in *Multitemporal Remote Sensing: Method and Applications*, Ed. Y.F. Ban, Springer Verlag, 2015. **Submitted.**

INTERNATIONAL JOURNALS

[J1] **S. Liu**, L. Bruzzone, F. Bovolo, P. Du., "Unsupervised Multitemporal Spectral Unmixing for Detecting Multiple Changes In Hyperspectral Images," *IEEE Transactions on Geoscience and Remote Sensing,* **under revision**.

[J2] **S. Liu**, L. Bruzzone, F. Bovolo, M. Zanetti, P. Du., "Sequential Spectral Change Vector Analysis for Iteratively Discovering and Detecting Multiple Changes in Hyperspectral Images", *IEEE Transactions on Geoscience and Remote Sensing,* **accepted for publication**.

[J3] **S. Liu**, L. Bruzzone, F. Bovolo, P. Du, "Hierarchical Unsupervised Change Detection in Multitemporal Hyperspectral Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol.53, no.1, pp: 244-260, 2015.

[J4] P. Du, P. Liu, J. Xia, F. Li, **S. Liu**, K. Tan, L. Cheng, "Remote Sensing Image Interpretation for Urban Environment Analysis: Methods, System and Examples," *Remote Sensing*, vol.6, no.10, pp: 9458-9474, 2014.

[J5] P. Du, **S. Liu**, P. Liu, K. Tan and L. Cheng, "Sub-pixel change detection for urban land-cover analysis via multi-temporal remote sensing images," *Geo-spatial Information Science*, vol.17, no.1, pp:26-38, 2014.

[J6] A. Samat, P. Du, **S. Liu**, J. Li, L. Cheng, "E^2LMs: Ensemble Extreme Learning Machines for Hyperspectral Image Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol.7, no.4, pp.1060-1069, 2014.

[J7] C. Yang, **S. Liu**, L. Bruzzone, R. Guan, P. Du, "A feature-metric-based affinity propagation technique for feature selection in hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol.10, no.5, pp.1152-1156, 2013.

[J8] P. Du, **S. Liu**, J. Xia, Y. Zhao, "Information fusion techniques for change detection from multi-temporal remote sensing images," *Information Fusion*, vol. 14, no.1, pp. 19-27, 2013.

[J9] P. Du, **S. Liu**, P. Gamba, K. Tan, J. Xia, "Fusion of difference images for change detection over urban areas," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol.5, no.4, pp.1076-1086, 2012.

INTERNATIONAL CONFERENCES

[C1] **S. Liu**, L. Bruzzone, F. Bovolo, P. Du, "Multitemporal Spectral Unmixing for Change Detection in Hyperspectral Images", *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, **submitted**.

[C2] J. Li, L. Bruzzone, **S. Liu**, “Deep Feature Representation for Hyperspectral Image Classification,” *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, **submitted**.

[C3] **S. Liu**, L. Bruzzone, F. Bovolo, P. Du, “A novel sequential spectral change vector analysis for representing and detecting multiple changes in hyperspectral images,” *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, vol., no., pp.4656-4659, 13-18 July 2014.

[C4] **S. Liu**, L. Bruzzone, F. Bovolo, P. Du, “A novel hierarchical method for change detection in multitemporal hyperspectral images,” *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, vol., no., pp.823-826, 21-26 July 2013.

[C5] P. Du, **S. Liu**, L. Bruzzone, F. Bovolo, “Target-driven change detection based on data transformation and similarity measures,” *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, vol., no., pp.2016, 2019, 22-27 July 2012.

[C6] **S. Liu**, L. Bruzzone, F. Bovolo, P. Du, “Unsupervised hierarchical spectral analysis for change detection in hyperspectral images,” *4th workshop on hyperspectral image and signal processing: evolution in remote sensing*. Shanghai, China, 4-7 June, 2012.

[C7] C. Yang, **S. Liu**, L. Bruzzone, R. Guan, P. Du, “A semisupervised feature metric based band selection method for hyperspectral image classification,” *4th workshop on hyperspectral image and signal processing: evolution in remote sensing*. Shanghai, China, 4-7 June, 2012.

Bibliography

[1] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Springer, 2006.

[2] J. M. Bioucas-Dias, et al., "Hyperspectral Remote Sensing Data Analysis and Future Challenges," *Geoscience and Remote Sensing Magazine, IEEE*, vol. 1, no. 2, pp. 6-36, 2013.

[3] R. Beck. (15 July 2003) EO-1 User Guide.

[4] C.-I. Change, *Hyperspectral Imaging: Signal Processing Algorithm Design and Analysis*. John Wiley & Sons, Inc., 2013.

[5] L. Egghe, "Untangling Herdan's law and Heaps' law: Mathematical and informetric arguments," *Journal of the American Society for Information Science and Technology*, vol. 58, no. 5, p. 702–709, 2007.

[6] L. Bruzzone and F. Bovolo, "A Novel Framework for the Design of Change-Detection Systems for Very-High-Resolution Remote Sensing Images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 609-630, 2013.

[7] D. Lu, P. Mause, E. Brondízio, and E. Moran, "Change detection techniques," *International Journal of Remote Sensing* , vol. 25, no. 12, pp. 2365-2401, 2004.

[8] M. K. Ridd and J. Liu, "A Comparison of Four Algorithms for Change Detection in an Urban Environment," *Remote Sensing of Environment*, vol. 63, no. 2, pp. 95-100, 1998.

[9] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: a review," *International Journal of Remote Sensing* , vol. 25, no. 9, pp. 1565-1596, 2004.

[10] D. Lu, G. Li, and E. Moran, "Current situation and needs of change detection techniques," *International Journal of Image and Data Fusion*, vol. 5, no. 1, pp. 13-38, 2014.

[11] P. R. Coppin and M. E. Bauer, "Digital change detection in forest ecosystems with remote sensing imagery," *Remote Sensing Reviews*, vol. 13, no. 3-4, pp. 207-304, 1996.

[12] P. Du, S. Liu, P. Gamba, K. Tan, and J. Xia, "Fusion of Difference Images for Change Detection over Urban Areas," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 4, pp. 1076-1086, 2012.

[13] P. Du, S. Liu, and K. Tan, "Rapid monitoring of japan earthquake-triggered tsunami disaster based on a fusion of multiple features derived from HJ small satellite images," *Science and Technology Review*, vol. 30, no. 4, pp. 31-36, 2012.

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

[15] J. Gong, H. Sui, G. Ma, and Q. Zhou, "A review of multi-temporal remote sensing data change detection algorithms," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, no. B7, pp. 757-762, 2008.

[16] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294-307, 2005.

[17] A. Singh, "Digital change detection techniques using remotely-sensed data," *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989-1003, 1989.

[18] L. Castellana, A. D'Addabbo, and G. Pasquariello, "A composed supervised/unsupervised approach to improve change detection from remote sensing," *Pattern Recogn. Lett.*, vol. 28, no. 4, pp. 405-413, 2007.

[19] X. L. Dai and S. Khorram, "Remotely sensed change detection based on artificial neural networks," *Photogrammetric engineering and remote sensing*, vol. 65, no. 10, pp. 1187-1194, 1999.

[20] X. Li and A. G. O. Yeh, "Principal component analysis of stacked multi-temporal images for the monitoring of rapid urban expansion in the Pearl River Delta," *International Journal of Remote*

Sensing, vol. 19, no. 8, pp. 1501-1518, 1998.

[21] H. Nemmour and Y. Chibani, "Multiple support vector machines for land cover change detection: An application for mapping urban extensions," *ISPRS J Photogramm*, vol. 61, no. 2, pp. 125-133, 2006.

[22] V. P. Soares and R. M. Hoffer, "Eucalyptus forest change classification using multi-data Landsat TM data," *Proceedings, SPIE*, vol. 2314, pp. 281-291, 1994.

[23] P. Du, S. Liu, and H. Zheng, "Land cover change detection over mining areas based on support vector machine," *Journal of China University of Mining and Technology*, vol. 41, no. 2, pp. 262-267, 2012.

[24] S. Liu, P. Du, P. Gamba, and a. J. Xia, "Fusion of difference images for change detection in urban areas," in *Urban Remote Sensing Event (JURSE)*, Munich, 2011.

[25] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. on Inf. Theory*, vol. IT-14, pp. 55-63, 1968.

[26] L. Bruzzone and R. Cossu, "A multiple cascade-classifier system for a robust a partially unsupervised updating of land-cover maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 1984-1996, 2002.

[27] B. Demir, F. Bovolo, and L. Bruzzone, "Detection of Land-Cover Transitions in Multitemporal Remote Sensing Images With Active-Learning-Based Compound Classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 5, pp. 1930-1941, 2012.

[28] O. Nobuyuki, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. IEEETransactionsonSystems,ManandCyberneticsInN/A,Vol9,No1(January1979),62-66,doi:101109/tsmc19794310076Key:citeulike:2917492.

[29] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, vol. 19, no. 1, pp. 41-47, 1986.

[30] P. K. Sahoo, S. Soltani, and A. K. C. Wong, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing*, vol. 41, no. 2, p. 233–260, 1988.

[31] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *Geoscience and Remote Sensing, IEEE Transactions on* , vol. 38, no. 3, pp. 1170-1182, 2000.

[32] L. Bruzzone and D. F. Prieto, "A minimum-cost thresholding technique for unsupervised change detection," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3539-3544, 2000.

[33] F. Bovolo and L. Bruzzone, "A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 1, pp. 218-236, 2007.

[34] T. Celik, "Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering," *Geoscience and Remote Sensing Letters, IEEE* , vol. 6, no. 4, pp. 772-776, 2009.

[35] A. Ghosh, N. S. Mishra, and S. Ghosh, "Fuzzy clustering algorithms for unsupervised change detection in remote sensing images," *Information Sciences*, vol. 181, no. 4, pp. 699-715, 2011.

[36] F. De Morsier, D. Tuia, V. Gass, J. Thiran, and M. Borgeaud, "Unsupervised change detection via hierarchical support vector clustering," in *Pattern Recognition in Remote Sensing (PRRS), 2012 IAPR Workshop on*, 2012.

[37] M. Gong, S. Linzhi, J. Meng, and W. Chen, "Fuzzy Clustering With a Modified MRF Energy Function for Change Detection in Synthetic Aperture Radar Images," *Fuzzy Systems, IEEE Transactions on*, vol. 22, no. 1, pp. 98-109, 2014.

[38] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised Change Detection With Kernels," *Geoscience and Remote Sensing Letters, IEEE*, vol. 9, no. 6, pp. 1026-1030, 2012.

[39] P. Gong, "Change detection using principal components analysis and fuzzy set theory," *Canadian Journal of Remote Sensing*, vol. 19, pp. 22-29, 1993.

[40] J. B. Collins and C. E. Woodcock, "Change detection using the Gramm-Schmidt transformation

applied to mapping forest mortality," *Remote Sensing of Environment*, vol. 50, no. 3, pp. 267-279, 1994.

[41] J. Rogan, J. Franklin, and D. A. Roberts, "A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery," *Remote Sensing of Environment*, vol. 80, no. 1, pp. 143-156, 2002.

[42] M. Frank and M. Canty, "Unsupervised Change Detection for Hyperspectral Images," in *In 12th JPL Airborne Earth Science Workshop*, 2003.

[43] A. A. Nielsen, "The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data," *IEEE Trans. on Image Proc.*, vol. 16, no. 2, pp. 463-478, 2007.

[44] W. A. Malila, "Change Vector Analysis: An Approach for. Detecting Forest Changes with Landsat," in *6th Annual Symposium on Machine Processsing of Remotely Sensed Data*, West Lafayette, Indiana, 1980, pp. 326-336.

[45] E. F. Lambin and A. H. Strahler, "Change-vector analysis in multitemporal space: A tool to detect and categorize land-cover change processes using high temporal-resolution satellite data," *Remote Sensing of Environment*, vol. 48, pp. 231-244, 1994.

[46] J. Chen, P. Gong, C. He, R. Pu, and P. Shi, "Land-use/land-cover change detection using improved change-vector analysis," *Photogrammetric Engineering & Remote Sensing*, vol. 69, no. 4, pp. 369-379, 2003.

[47] F. Bovolo, S. Marchesi, and L. Bruzzone, "A Framework for Automatic and Unsupervised Detection of Multiple Changes in Multitemporal Images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 6, pp. 2196-2212, 2012.

[48] F. Bovolo and L. Bruzzone, "An adaptive thresholding approach to multiple-change detection in multispectral images," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, Vancouver, BC, 2011, pp. 233-236.

[49] R. Wiemker, A. Speck, D. Kulbach, H. Spitzer, and J. Bienlein, "Unsupervised Robust Change Detection on Multispectral Imagery Using Spectral and Spatial Features," in *Proceedings of the Third International Airborne Remote Sensing Conference and Exhibition*, 1997, pp. 640-647.

[50] G. Camps-Valls, et al., "Multitemporal image classification and change detection with kernels," in *Remote Sensing. International Society for Optics and Photonics*, 2006.

[51] N. Falco, M. Dalla Mura, F. Bolovo, J. A. Benediktsson, and L. Bruzzone, "Change detection in VHR images based on morphological attribute profiles," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 3, pp. 636-640, 2013.

[52] P. Du and S. Liu, "Change detection from multi-temporal remote sensing images by integrating multiple features," *Journal of Remote Sensing*, vol. 16, no. 4, pp. 663-677, 2012.

[53] 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*, vol. 29, no. 2, pp. 399-423, 2008.

[54] B. Desclée, P. Bogaert, and P. Defourny, "Forest change detection by statistical object-based method," *Remote Sensing of Environment*, vol. 102, pp. 1-11, 2006.

[55] F. Bovolo, "A Multilevel Parcel-Based Approach to Change Detection in Very High Resolution Multitemporal Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 1, pp. 33-37, 2009.

[56] S. Liu and P. Du, "Object-oriented change detection from multi-temporal remotely sensed images," in *GEOBIA 2010: Geographic Object-Based Image Analysis*, Ghent, Belgium, 2010, pp. VolXXXVIII-4/C7.

[57] T. Celik, "Change detection in satellite images using a genetic algorithm approach," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 386-390, 2010.

[58] P. Du, S. Liu, J. Xia, and Y. Zhao, "Information fusion techniques for change detection from multi-temporal remote sensing images," *Information Fusion*, vol. 14, no. 1, pp. 19-27, 2013.

[59] C. Wu, B. Du, and L. Zhang, "Slow Feature Analysis for Change Detection in Multispectral

Imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 5, pp. 2858-2874, 2014.

[60] H.-C. Li, N. Longbotham, and W. L. Emery, "Unsupervised change detection of remote sensing images based on semi-nonnegative matrix factorization," in *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, 2014, pp. 1289-1292.

[61] K. Ding, C. Huo, Y. Xu, Z. Zhong, and C. Pan, "Sparse Hierarchical Clustering for VHR Image Change Detection," *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, vol. 12, no. 3, pp. 577-581, 2015.

[62] A. Schaum and A. Stocker, "Long-interval chronochrome target detection," in *Int. Symp. Spectral Sens. Res.*, 1997.

[63] A. Schaum and A. Stocker, "Hyperspectral change detection and supervised matched filtering based on covariance equalization," in *Proc. SPIE*, vol. 5425, 2004, pp. 77-90.

[64] M. Shimoni, R. Heremans, and C. ,. Perneel, "Detection of small changes in complex urban and industrial scenes using imaging spectroscopy," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, Honolulu, HI , 2010, pp. 4757-4760.

[65] A. A. Nielsen, "Kernel Maximum Autocorrelation Factor and Minimum Noise Fraction Transformations," *Image Processing, IEEE Transactions on* , vol. 20, no. 3, pp. 612-624, 2011.

[66] V. Ortiz-Rivera, M. Vélez-Reyes, and B. Roysam, "Change detection in hyperspectral imagery using temporal principal components," in *Proc. SPIE 6233,Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII*, Orlando (Kissimmee), FL, 2006, p. 623312.

[67] A. A. Nielsen, R. Larsen, and J. S. Vestergaard, "Sparse principal component analysis in hyperspectral change detection," in *SPIE Europe Remote Sensing Conference 8180*, Prague, Czech Republic, 2011.

[68] S. Liu, L. Bruzzone, F. Bovolo, and P. Du, "Unsupervised hierarchical spectral analysis for change detection in hyperspectral images," in *Hyperspectral Image and Signal Processing (WHISPERS), 2012 4th Workshop on* , Shanghai, 2012, pp. 1-4.

[69] S. Himissi, K. Saheb Ettabaa, I. R. Farah, and B. Soulaiman, "A New Spatio-temporal ICA for Multi-temporal Endmembers Extraction and Change Trajectory Analysis," in *PIERS Proceedings: Progress In Electromagnetics Research Symposium*, 2011, pp. 1289-1293.

[70] S. Adar, G. Notesco, A. Brook, and e. a. Livne, "Change detection over Sokolov open-pit mining area, Czech Republic, using multi-temporal HyMAP data (2009-2010)," in *Proc. SPIE 8180, Image and Signal Processing for Remote Sensing XVII, 81800T.*, 2011.

[71] Y. Z. Du, C. I. Chang, H. Ren, C. C. Chang, and J. O. Jensen, "New hyperspectral discrimination measure for spectral characterization," *Opt. Eng.*, vol. 43, no. 8, pp. 1777-1786, 2004.

[72] K. Vongsy and M. J. Mendenhall, "A comparative study of spectral detectors," in *Proc. 3rd WHISPERS*, Lisbon, Portugal, 2011.

[73] Q. Du, L. Wasson, and R. King, "Unsupervised linear unmixing for change detection in multitemporal airborne hyperspectral imagery," in *Int. Workshop on MultiTemp*, 2005, pp. 136-140.

[74] Q. Du, N. Younan, and R. King, "Change Analysis for Hyperspectral Imagery," in *Int. Workshop on MultiTemp*, 2007, pp. 1-4.

[75] S. Hemissi, I. R. Farah, K. Saheb Ettabaa, and B. Solaiman, "Multi-Spectro-Temporal Analysis of Hyperspectral Imagery Based on 3-D Spectral Modeling and Multilinear Algebra," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 1, pp. 199-216, 2013.

[76] C. Wu, B. Du, and L. Zhang, "A Subspace-Based Change Detection Method for Hyperspectral Images," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 6, no. 2, pp. 815-830, 2013.

[77] Q. Du, "A new method for change analysis of multi-temporal hyperspectral images," in *Proc. 4td WHISPERS*, 2012.

[78] J. Meola, M. T. Eismann, R. L. Moses, and J. N. Ash, "Detecting Changes in Hyperspectral Imagery Using a Model-Based Approach," *IEEE Trans. Geosci. Rem. Sens.*, vol. 49, no. 7, pp.

2647-2661, 2011.

[79] J. Meola, M. T. Eismann, R. L. Moses, and J. N. Ash, "Application of Model-Based Change Detection to Airborne VNIR/SWIR Hyperspectral Imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 10, pp. 3693-3706, 2012.

[80] J. Yang and W. Sun, "Slight Change Extraction and Detection in Multitemporal Hyperspctral Images," *Journal of Computational Information Systems*, vol. 10, no. 15, pp. 6703-6711, 2014.

[81] Y. Yuan, H. L. Lv, and X. Lu, "Semi-supervised change detection method for multi-temporal hyperspectral images," *Neurocomputing*, vol. 148, no. 19, pp. 363-375, 2015.

[82] K. Vongsy, M. J. Mendenhall, M. T. Eismann, and G. L. Peterson, "Removing parallax-induced changes in Hyperspectral Change Detection," in *IEEE Int. Geosci. and Remote Sens. Symp.*, 2012.

[83] J. Meola, M. T. Eismann, K. J. Barnard, and R. C. Hardie, "Analysis of hyperspectral change detection as affected by vegetation and illumination variation," in *Proc. SPIE*, 2007.

[84] M. T. Eismann, J. Meola, A. Stocker, S. Beaven, and A. Schaum, "Hyperspectral change detection in the presence of diurnal and seasonal variations," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 237-249, 2008.

[85] M. Dalponte, L. Bruzzone, L. Vescovo, and D. Gianelle, "The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas," *Remote Sensing of Environment*, vol. 113, pp. 2345-2355, 2009.

[86] S. Liu, L. Bruzzone, F. Bovolo, and P. Du, "Hierarchical change detection in multitemporal hyperspectral images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, no. 1, pp. 244-260, 2015.

[87] S. Liu, L. Bruzzone, F. Bovolo, and P. Du, "A novel sequential spectral change vector analysis for representing and detecting multiple changes in hyperspectral images," in *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, 2014, pp. 4656-4659.

[88] F. Bovolo, S. Marchesi, and L. Bruzzone, "A Framework for Automatic and Unsupervised Detection of Multiple Changes in Multitemporal Images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 6, pp. 2196-2212, 2012.

[89] T. Fung, "An assessment of TM imagery for land cover change detection," *Transactions on Geoscience and Remote Sensing, IEEE*, vol. 28, no. 4, pp. 681-684, 1990.

[90] J. B. Collins and C. E. Woodcock, "An assessment of several linear change detection techniques for mapping forest mortality using multitemporal landsat TM data," *Remote Sensing of Environment*, vol. 56, no. 1, pp. 66-77, 1996.

[91] J. Collins and C. Woodcock, "Change detection using the Gramm-Schmidt transformation applied to mapping forest mortality," *Remote Sensing of Environment*, vol. 50, no. 3, pp. 267-279, 1994.

[92] L. Bruzzone and S. B. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 35, no. 4, pp. 858-867, 1997.

[93] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. New York: Plenum Publishing Co., 2003.

[94] N. Keshava, "Distance Metrics and Band Selection in Hyperspectral Processing With Applications to Material Identification and Spectral Libraries," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 7, pp. 1552-1565, 2004.

[95] C. I. Chang, "An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image," *IEEE Transactions on Information Theory*, vol. 46, no. 5, pp. 1927-1932, 2000.

[96] Y. Sohn and N. S. Rebello, "Supervised and unsupervised spectral angle classifiers," *Photogrammetric Engineering and Remote Sensing*, vol. 68, no. 12, pp. 1271-1280, 2002.

[97] P. Du, S. Liu, L. Bruzzone, and F. Bovolo, "Target-driven change detection based on data transformation and similarity measures," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, Munich, Germany, 2012, pp. 2016-2019.

[98] K. L. Taylor. (2014) Oracle® Data Mining Concepts.

[99] J. Li. (2011) Discriminative Image Segmentation: Applications to Hyperspectral Data. Ph.D. Dissertation .

[100] Hyperspectral Remote Sensing Scenes. Computational Intelligence Group from the Basque University (UPV/EHU).. [Online]. http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

[101] A. Chakrabarti and T. Zickler, "Statistics of Real-World Hyperspectral Images," in *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 193-200.

[102] Earth science data archives of U.S. Geological Survey (USGS). [Online]. http://earthexplorer.usgs.gov/

[103] Bishop and C.M., *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag,, 2006.

[104] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650-663, 2004.

[105] D. Pelleg and A. W. Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters," in *In Proceedings of the Seventeenth International Conference on Machine Learning (ICML '00)*, San Francisco, CA, USA, 20000.

[106] T. Ishioka, "An expansion of x-means for automatically determining the optimal number of clusters - Progressive iterations of k-means and merging of the clusters," in *In: IASTED International Conference on Computational Intelligence*, 2005.

[107] X. Hu and L. Xu, "Investigation on several model selection criteria for determining the number of cluster," in *Neural Inform. Proces. Lett. and Reviews*, vol. 4, 2004, p. 2004.

[108] S. Tu and L. Xu, "A theoretical investigation of several model selection criteria for dimensionality reduction," *Pattern Recognition Letters*, vol. 33, no. 9, pp. 1117-1126, 2012.

[109] N. Keshava and J. F. Mustard, "Spectral unmixing," *Signal Processing Magazine, IEEE* , vol. 19, no. 1, pp. 44-57, 2002.

[110] N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55-78, 2003.

[111] J. M. Bioucas-Dias, et al., "Hyperspectral Remote Sensing Data Analysis and Future Challenges," *Geoscience and Remote Sensing Magazine, IEEE*, vol. 1, no. 2, pp. 6-36, 2013.

[112] D. B. Lobell and G. P. Asner, "Cropland distributions from temporal unmixing of MODIS data," *Remote Sensing of Environment*, vol. 93, no. 3, pp. 412-422, 2004.

[113] B. Somers and G. P. Asner, "Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests," *Remote Sensing of Environment*, vol. 136, pp. 14-27, 2013.

[114] M. A. Goenaga, M. C. Torres-Madronero, M. Velez-Reyes, S. J. Van Bloem, and J. D. Chinea, "Unmixing Analysis of a Time Series of Hyperion Images Over the Guánica Dry Forest in Puerto Rico," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* , vol. 6, no. 2, pp. 329-338, 2013.

[115] P. Du, S. Liu, P. Liu, K. Tan, and L. Cheng, "Sub-pixel change detection for urban land-cover analysis via multi-temporal remote sensing images," *Geo-spatial Information Science*, vol. 17, no. 1, pp. 26-38, 2014.

[116] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in Spectral Mixture Analysis: A review," *Remote Sensing of Environment*, vol. 115, no. 7, pp. 1603-1616, 2011.

[117] A. Zare and K. C. Ho, "Endmember Variability in Hyperspectral Analysis: Addressing Spectral Variability During Spectral Unmixing," *Signal Processing Magazine, IEEE* , vol. 31, no. 1, pp. 95-104, 2014.

[118] J. Harsanyi, W. Farrand, and C. .-I. Chang, "Determining the number and identity of spectral endmembers; an integrated approach using Neyman-Person eigen-thresholding and iterative constrained RMS error minimization," in *Proc. 9th Thematic Conf. Geologic Remote Sensing*,

1993.

[119] C.-I. Chang and Q. D. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 3, pp. 608-619, 2004.

[120] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435-2445, 2008.

[121] B. Luo, J. Chanussot, S. Doute, and L. Zhang, "Empirical Automatic Estimation of the Number of Endmembers in Hyperspectral Images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 1, pp. 24-28, 2013.

[122] J. Nascimento and J. Bioucas-Dias, "Vertex Component Analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898-910, 2005.

[123] Y. Du, et al., "New hyperspectral discrimination measure for spectral characterization," *Optical Engineering*, vol. 43, no. 8, pp. 1777-1786, 2004.

[124] Google. Google Maps. [Online]. https://www.google.com.hk/maps/

[125] S. Marchesi and L. Bruzzone, "A registration-noise driven technique for the alignment of VHR remote sensing images," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, Honolulu, HI, 2010, pp. 1023-1026.

Appendix A

In this appendix, the mathematical definitions of Euclidean distance Δ and Spectral angle distance ϑ two metrics are recalled. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two vectors in a B-dimensional space. The two distance metrics are defined as follows:

$$\Delta(\boldsymbol{X},\boldsymbol{Y}) = \|\boldsymbol{X} - \boldsymbol{Y}\| = \sqrt{\sum_{b=1}^{B}(x_b - y_b)} \tag{35}$$

$$\vartheta(\boldsymbol{X},\boldsymbol{Y}) = \arccos\left(\frac{\langle \boldsymbol{X},\boldsymbol{Y}\rangle}{\|\boldsymbol{X}\|\|\boldsymbol{Y}\|}\right) = \arccos\left(\sum_{b=1}^{B} x_b y_b \Big/ \sqrt{\sum_{b=1}^{B}(x_b)^2 \sum_{b=1}^{B}(y_b)^2}\right) \tag{36}$$

TABLE A-1 PROPERTIES OF EUCLIDEAN DISTANCE AND SPECTRAL ANGLE DISTANCE **[94]**

Metric	**Euclidean distance**	**Spectral angle distance**
Value	$0 \le \Delta$	$0 \le \vartheta \le \frac{\pi}{2}$
Invariance	Rotational	Multiplicative scaling
Additivity	Yes	No
Monotonicity	Yes	No

The properties of two distance metrics are summarized in TABLE A-1. In particular, it is important to note that spectral angle distance is invariant to the multiplicative scaling (37). Therefore, according to this property, the reference vector $\boldsymbol{R}$ for computing α in C^2VA is not only limited to a unit vector $\boldsymbol{R} = [1/\sqrt{B},\ldots,1/\sqrt{B}]$, it can be extended to any constant vector $\boldsymbol{R} = [a,\ldots,a]$, where $a > 0$ and $a \in \Re$.

$$\begin{aligned}\vartheta(a\boldsymbol{X},b\boldsymbol{Y}) &= \arccos\left(\frac{\langle a\boldsymbol{X},b\boldsymbol{Y}\rangle}{\|a\boldsymbol{X}\|\|b\boldsymbol{Y}\|}\right)\\ &= \arccos\left(\frac{ab\langle \boldsymbol{X},\boldsymbol{Y}\rangle}{ab\|\boldsymbol{X}\|\|\boldsymbol{Y}\|}\right)\\ &= \vartheta(\boldsymbol{X},\boldsymbol{Y}).\end{aligned} \tag{37}$$

where a, $b \in \Re$.

However, the selection of a specific vector $\boldsymbol{R}$ that points out the major variations present in the high-dimensional feature space (i.e., in our case are the multiple changes) is the key issue, which can result in a reliable compressed change representation in a 2-Dimensional polar domain.

Appendix B

As an alternative technique, Singular Value Decomposition (SVD) is also suitable to help extract a reference vector $\boldsymbol{R}$, which used for building a reasonable and effective representation of the hidden change patterns in Ω. From the mathematical point of view, SVD is a factorization approach that decomposes the matrix $\boldsymbol{x}_{h,j}$ (i.e., SCVs in the considered cluster $P_{h,j}$, cluster j at level h of the hierarchy) into a product of three matrices:

$$\boldsymbol{x}_{h,j} = \boldsymbol{U}^{s}_{h,j}\boldsymbol{D}^{s}_{h,j}\left(\boldsymbol{V}^{s}_{h,j}\right)^{*} \tag{38}$$

where two unitary matrices $\boldsymbol{U}^{s}_{h,j}$ and $\boldsymbol{V}^{s}_{h,j}$ represent sets of 'left' and 'right' orthonormal bases, respectively. $\left(\boldsymbol{V}^{s}_{h,j}\right)^{*}$ denotes the conjugate transpose of the unitary matrix $\boldsymbol{V}^{s}_{h,j}$. $\boldsymbol{D}^{s}_{h,j}$ is a diagonal matrix where the singular values are sorted in a descending order (i.e., $\lambda^{s,1}_{h,j} > \lambda^{s,2}_{h,j} > ...$) in the diagonal:

$$\boldsymbol{D}^{s}_{h,j} = \begin{bmatrix} \lambda^{s,1}_{h,j} & & \cdots & 0 \\ & \lambda^{s,2}_{h,j} & & \vdots \\ \vdots & & \lambda^{s,3}_{h,j} & \\ 0 & \cdots & & \ddots \end{bmatrix} \tag{39}$$

The singular values are all non-negative. Their magnitudes indicate the importance of the corresponding bases (i.e., vectors in $\boldsymbol{U}^{s}_{h,j}$ and $\boldsymbol{V}^{s}_{h,j}$). Thus singular values can reflect the amount of data variance captured by the bases. In particular, the first vector $\boldsymbol{V}^{s,1}_{h,j}$ (corresponding to the biggest singular value $\lambda^{s,1}_{h,j}$) in matrix $\boldsymbol{V}^{s}_{h,j}$ shows a reference direction that maximizes the variance of the measurement on $\alpha_{h,j}$.

$$\boldsymbol{R}_{h,j} = \boldsymbol{V}^{s,1}_{h,j} = \begin{pmatrix} v^{s}_{1} \\ v^{s}_{2} \\ \vdots \\ v^{s}_{B} \end{pmatrix}_{h,j} \tag{40}$$

The SVD in general has a more generalize ability due to the fact that it can be applied to non-square matrix, whereas the eigenvalue decomposition can be only applied to certain classes of square matrices. However, two decompositions are still related. Given a SVD of $\boldsymbol{x}_{h,j}$ (see (38)), the following two relations hold:

$$\boldsymbol{x}^{*}_{h,j}\boldsymbol{x}_{h,j} = \boldsymbol{V}^{s}_{h,j}\left(\boldsymbol{D}^{s}_{h,j}\right)^{*}\left(\boldsymbol{U}^{s}_{h,j}\right)^{*}\boldsymbol{U}^{s}_{h,j}\boldsymbol{D}^{s}_{h,j}\left(\boldsymbol{V}^{s}_{h,j}\right)^{*} = \boldsymbol{V}^{s}_{h,j}\left[\left(\boldsymbol{D}^{s}_{h,j}\right)^{*}\boldsymbol{D}^{s}_{h,j}\right]\left(\boldsymbol{V}^{s}_{h,j}\right)^{*} \tag{41}$$

$$\boldsymbol{x}_{h,j}\boldsymbol{x}_{h,j}^{*} = \boldsymbol{U}_{h,j}^{s}\boldsymbol{D}_{h,j}^{s}\left(\boldsymbol{V}_{h,j}^{s}\right)^{*}\boldsymbol{V}_{h,j}^{s}\left(\boldsymbol{D}_{h,j}^{s}\right)^{*}\left(\boldsymbol{U}_{h,j}^{s}\right)^{*} = \boldsymbol{U}_{h,j}^{s}\left[\boldsymbol{D}_{h,j}^{s}\left(\boldsymbol{D}_{h,j}^{s}\right)^{*}\right]\left(\boldsymbol{U}_{h,j}^{s}\right)^{*} \tag{42}$$

The right-hand formulations describe the eigenvalue decomposition of the left-hand ones, thus one can observe that:

1) Columns in $\boldsymbol{V}_{h,j}^{s}$ (right-singular vectors) are eigenvectors of $\boldsymbol{x}_{h,j}^{*}\boldsymbol{x}_{h,j}$;
2) Columns in $\boldsymbol{U}_{h,j}^{s}$ (left-singular vectors) are eigenvectors of $\boldsymbol{x}_{h,j}\boldsymbol{x}_{h,j}^{*}$;
3) Non-zero elements of $\boldsymbol{D}_{h,j}^{s}$ (non-zero singular values) are the square roots of the non-zero eigenvalues of $\boldsymbol{x}_{h,j}^{*}\boldsymbol{x}_{h,j}$ or $\boldsymbol{x}_{h,j}\boldsymbol{x}_{h,j}^{*}$.