Title: Unsupervised Multitemporal Spectral Unmixing for Detecting Multiple Changes in

Hyperspectral Images

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

Date of Publication: 12 January 2016

Author(s): S. Liu, L. Bruzzone, F. Bovolo, P. Du

Volume: 54, Issue: 5

Page(s): 2733 - 2748

DOI: 10.1109/TGRS.2015.2505183

UNSUPERVISED MULTITEMPORAL SPECTRAL UNMIXING FOR DETECTING MULTIPLE CHANGES IN HYPERSPECTRAL IMAGES

Sicong Liu, *Student Member, IEEE*, Lorenzo Bruzzone, *Fellow, IEEE*, Francesca Bovolo, *Senior Member, IEEE*, Peijun Du, *Senior Member, IEEE*
e-mail: bruzzone@ing.unitn.it

ABSTRACT

This paper presents a novel 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 a pixel level, the proposed technique investigates the spectral-temporal variations at a subpixel level. The considered Change Detection (CD) problem is analyzed in a multitemporal domain, where a bitemporal spectral mixture model is defined to analyze the spectral composition within a pixel. Distinct *multitemporal endmembers* (MT-EMs) are extracted according to an automatic and unsupervised 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 analyzing 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.

Keywords

Change detection; multiple changes; spectral unmixing; hyperspectral images; multi-temporal images; land-cover transitions; unsupervised analysis; remote sensing.

I. INTRODUCTION

The new generation of hyperspectral (HS) satellite sensors can acquire images having a very high spectral resolution (*e.g.*, 5nm-10nm) over a wide wavelength spectrum (*e.g.*, 400nm-2500nm). Thus for each pixel a near-continuous spectral signature can be obtained over the whole range of wavelengths. This important property makes the observation and identification of more detailed and specific land-cover materials from satellite images possible. Traditional change-detection (CD) techniques are mainly developed for multispectral (MS) images and aim at automatically detecting the land-cover changes occurred between two (or more) images acquired over the same geographical area at different observation times [1]. However, due to the limited spectral resolution of the MS images, usually only strong changes that result in a significant change of the pixel spectral signatures are detected. With the increasing availability of multitemporal HS images, change detection in hyperspectral images (CD-HS) becomes a very interesting and valuable topic, which makes it possible a very detailed monitoring of land-cover dynamics.

By taking advantage of the fine spectral information in HS images, subtle changes (which are not visible when employing MS 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 HS images to address new applications. However, due to the intrinsic properties of HS data, this task is highly challenging [2]. Examples of such properties are the high-dimensionality of the feature space, the information redundancy, the noise and the presence of many possible changes.

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. In general, depending on the availability of the reference samples, two main groups of techniques exist to address the multiple-change detection task

[2]: supervised and unsupervised. We focus on the unsupervised CD methods that do not require any reference data or prior knowledge, thus being more attractive and important in the real CD-HS applications. In [3], an unsupervised Multivariate Alteration Detection (MAD) technique was used to highlight the multiple change information based on the Canonical Correlation Analysis (CCA) thus to detect the seasonal vegetation changes in multitemporal HS images. An improved version of this technique, named Iterative Reweighted MAD (IR-MAD), was proposed in [4]. IR-MAD provides more reliable output components and better emphasizes the detected changes. However, it requires a strong interaction with the end-users to select the most informative components that represent the specific changes of interest, which is always time consuming. Therefore, it is complex to use it for detecting all possible change classes, especially when the number of changes is large. Recently, the authors investigated the CD-HS problem from the point of view of spectral signature changes, and analyzed the hierarchical nature of changes [2]. Based on this analysis, a sequential change representation, discovery and detection approach was proposed in [5]. It allows users to identify kinds of changes and to implement an interactive change identification scheme according to a sequence of 2-D change representation scattergrams. A hierarchical clustering approach to detect different levels of spectral changes following a systematic top-down structure was proposed in [2]. In the experiments both techniques outperformed the state-of-the-art methods in addressing the challenging multiple-change detection problem in HS images.

Despite the usefulness of the aforementioned 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 HS images, mixed pixels are a common phenomenon that occurs in most of the cases. This phenomenon consists in a mixture of the light scattering of more than one distinct substance located in the area on the ground covered by one pixel [6], [7], [8]. This mixture is usually caused either by the limited spatial resolution of the sensors that includes different targets in a single pixel or by the combination of the distinct materials into a homogeneous mixture [6], [7]. 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 [6], [7], [8], [9], 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 HS 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 has been only partially investigated to address the endmembers variability issue in order to increase the representativeness of the extracted endmembers, thus improving either land-cover classification or specific change monitoring (*e.g.*, cropland, invasive species, forest, etc.) [10], [11], [12]. In the context of unmixing, methods have been proposed for change detection that rely on post-unmixing comparison or posterior probability comparison. They have the capability to investigate subpixel changes. Du et al. [13] proposed a linear mixture model for analyzing endmembers and abundances estimated from each single time image to address a binary CD problem.

Lu et al. developed a CD method based on the linear spectral mixture analysis (LSMA) of the multitemporal Landsat TM images and the analysis of the obtained differencing fraction images [14]. Recently, a subpixel level CD approach was developed to investigate the multiple composition evidence within pixels, thus to increase the binary CD accuracy [15]. Chen et al. designed a posterior probability comparison by using change vector analysis to reduce the effect of the cumulative error in post-classification comparison (PCC) [16]. However, these methods are all designed in a supervised framework relying on the availability of training samples. None of them addresses the challenging multiple-change detection problem in multitemporal HS images from the spectral unmixing point of view with unsupervised techniques.

In this paper, we propose a novel CD approach that is suitable and effective for detecting multiple change classes in HS 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 *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 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 CD map based on the abundance combination. The proposed MSU approach is validated on both simulated and real bi-temporal HS images. The experimental results confirm the effectiveness of the proposed method in performing multiple-change detection.

The rest of the paper is organized as follows. Section II presents the defined multitemporal unmixing model and the considered multiple-change detection problem in HS images, pointing out the main properties of the multitemporal domain and their role in CD. The proposed MSU method is illustrated in Section III. Section IV describes the HS data sets, and analyzes and discusses the experimental results. Section V draws the conclusion of this work.

II. MULTITEMPORAL SPECTRAL UNMIXING IN CHANGE DETECTION ON HYPERSPECTRAL IMAGES

Let $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ be two co-registered HS images having a size of $I \times J$ acquired over the same geographical area at 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) $(1 \leq i \leq I, 1 \leq j \leq J)$ in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively. B is the number of spectral channels of the considered 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.1). Thus only change or no-change case 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.1), 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.1, 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 behavior (see Fig.1). 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 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 mixture phenomenon and its effects on CD. The proposed change formulation and representation is based on this spectral mixture analysis.

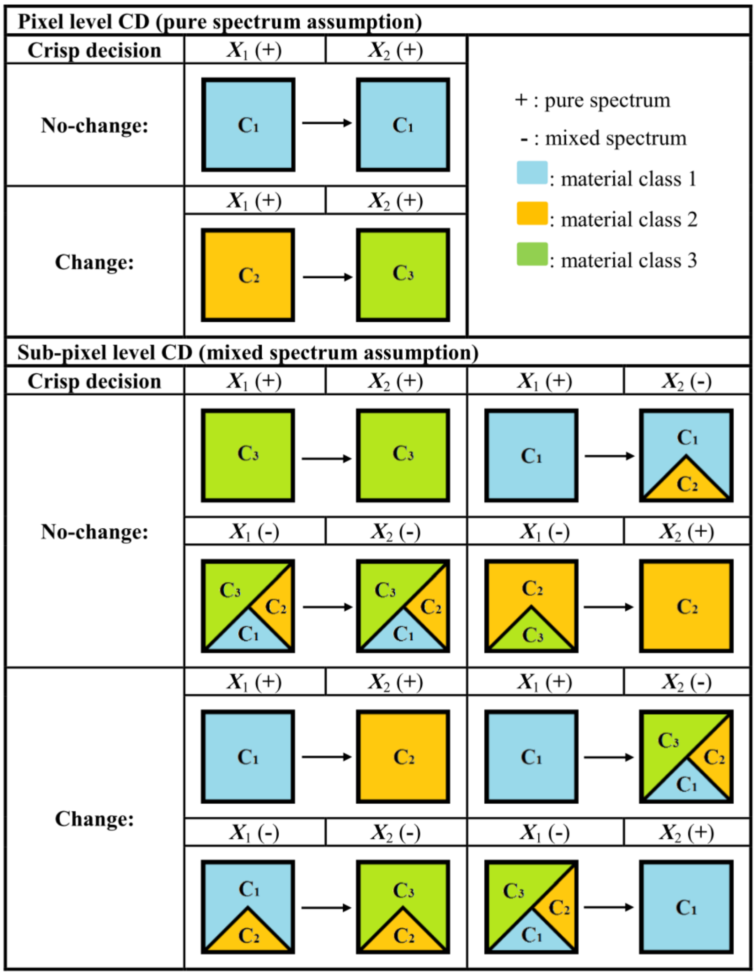


Fig.1 Possible change situations of a single pixel in the bi-temporal images based on the pure spectrum and mixed spectrum assumptions.

Sub-pixel (or pixel) transitions have been often detected by performing post-unmixing (or classification) comparison [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]. The difference between abundances (or class label) is used for identifying sub-pixel (or pixel) land-cover transitions. The performance of this kind of approaches strongly depends on the accuracy of each unmixing/classification process and on their correlation. However, when dealing with post-classification comparison, under the assumption of independent errors in single date classification, it is possible to stay that the accuracy comes close to the product of the accuracies yielded at the two times [17], [18], [21]. In unmixing the situation is more complex and the estimation of the cumulative post-unmixing comparison is still an open issue. The post-classification/unmixing approaches do not consider the temporal correlation between available acquisitions. The drawback becomes even more critical when dealing with hyperspectral images and in an unsupervised context.

A possible way to consider temporal correlation is to consider a change index. As we are dealing with optical passive sensors, the most popular index is the B-Dimensional difference image $\boldsymbol{X}_D$ computed by subtracting the multitemporal images pixel-by-pixel [2], [5], [24], [25], *i.e.*,

$$\boldsymbol{X}_D = \boldsymbol{X}_2 - \boldsymbol{X}_1 \tag{1}$$

The physical meaning of the pixel spectrum in $\boldsymbol{X}_D$ is associated to the land-cover transitions rather than to 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 order to preserve spectral signature information of both multitemporal images (as in post-unmixing comparison) and take advantage of the temporal correlation (as in CD in the spectral difference domain), we analyze the multitemporal unmixing problem in a $2B$-Dimensional multitemporal domain represented by $\boldsymbol{X}_S$, which is a stacked feature space based on the considered multitemporal images, *i.e.*,

$$\boldsymbol{X}_S = \left[\boldsymbol{X}_1, \boldsymbol{X}_2\right] \tag{2}$$

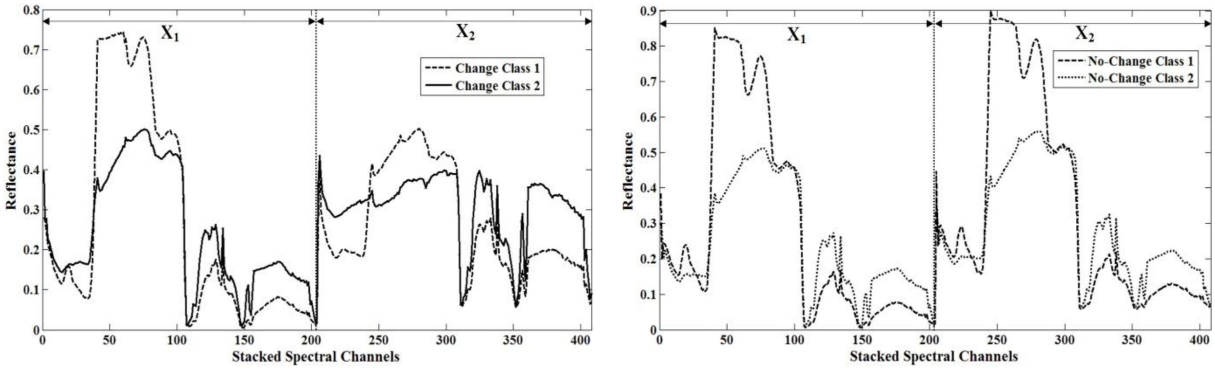
The multitemporal domain $\boldsymbol{X}_S$ has been used for CD purposes in the literature. Two main approaches can be identified: the supervised Direct Multi-date Classification (DMC) [26], [27] and the unsupervised stacked feature transformation [28], [29], [30], [31]. The former identifies changes by simultaneously classifying the stacked multi-date images, thus a land-cover transition is represented by an output class in the final classification map. The considered CD task is handled as a supervised classification task. However, the generation of a comprehensive training set that represents all the possible temporal land-cover transitions makes the supervised approach difficult to use in real applications. Sub-optimal solution to the classification problem can be implemented by using compound classification strategies [32]. The latter is implemented based on the multi-date transformation (*e.g.*, Temporal-Principal Components Analysis (T-PCA) [28], [29], Multi-date Kauth-Thomas (MKT) [30], Multi-date Graham-Schimidt (MGS) [31]), 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, in this paper we use a spectral mixture model in the multitemporal domain $\boldsymbol{X}_S$ to solve the considered multiple-change detection problem. The main advantages of working in such domain 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 $\boldsymbol{X}_S$ 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 $\boldsymbol{X}_S$ from two independent mixtures in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. Under the simplified assumption of linear mixture model between the two images, $\boldsymbol{X}_S$ can be described as:

$$\boldsymbol{X}_S = \left[\boldsymbol{X}_1, \boldsymbol{X}_2\right] = \left[\sum \boldsymbol{A}_1\boldsymbol{E}_1 + \boldsymbol{N}_1, \sum \boldsymbol{A}_2\boldsymbol{E}_2 + \boldsymbol{N}_2\right] \approx \sum \boldsymbol{A}_U\boldsymbol{U} + \boldsymbol{N}_S \tag{3}$$

where $\boldsymbol{U}$ is the matrix of the multitemporal endmember set, $\boldsymbol{A}_U$ is the corresponding abundance matrix, and $\boldsymbol{N}_S$ represents the noise matrix. $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, and $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ 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.

Illustrative examples of the spectral signatures of stacked channels associated with two change classes and two no-change classes in the $\boldsymbol{X}_S$ domain are shown in Fig.2 (a) and (b), respectively. The spectral signatures of the change classes have different spectral shapes in two components associated to $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ (see Fig.2.a), whereas the two components of the spectral signatures are almost the same for the no-change classes (see Fig.2.b). Moreover, different change and no-change classes have distinct spectral signatures in the multitemporal domain of $\boldsymbol{X}_S$ (see Fig.2 a and b).



(a) (b)

Fig.2 Examples of spectral signatures in the multitemporal domain X_S: (a) change and (b) no-change classes.

III. PROPOSED MULTITEMPORAL SPECTRAL UNMIXING CHANGE DETECTION APPROACH

In this paper, the considered sub-pixel CD problem is formalized as to estimate the abundance of change and no-change classes within a single pixel in $\boldsymbol{X}_S$. 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.3. 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. Details on each step are given in the following sub-sections.

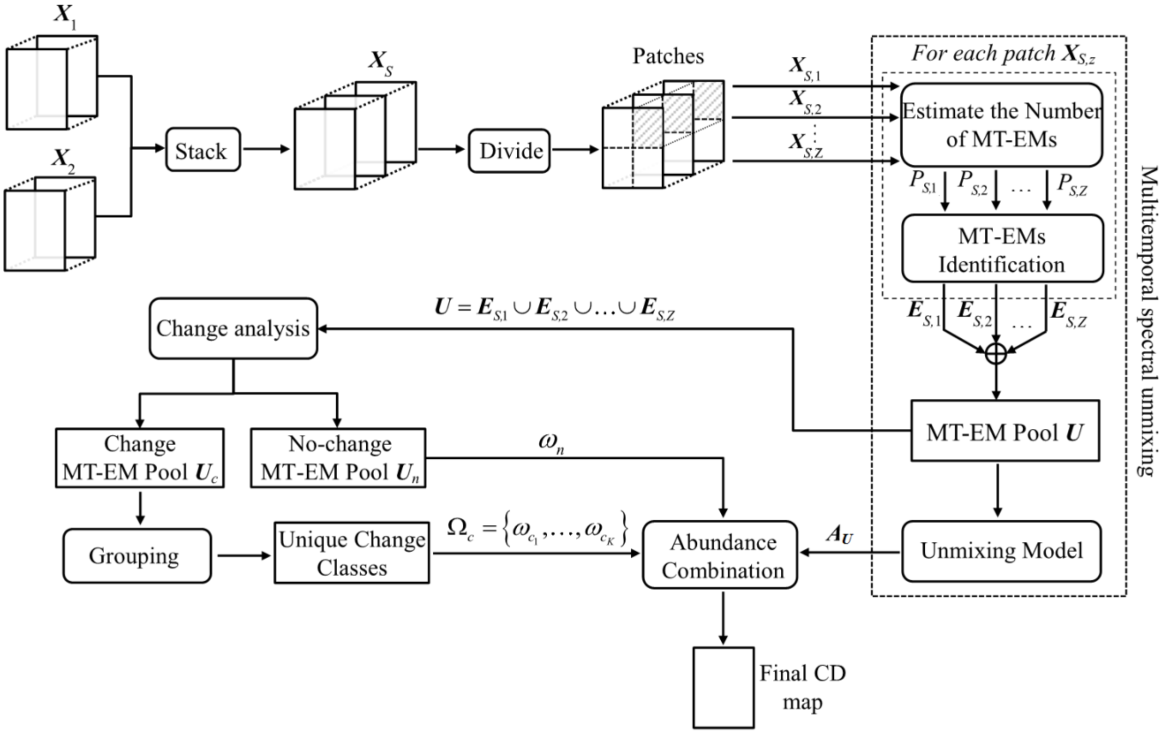


Fig.3 Architecture of the proposed CD approach based on multitemporal spectral unmixing.

A. Stacking of multitemporal images and image patches generation

In this step, the two B-Dimensional HS 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.2). Let P_1 and P_2 be the number of endmembers in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. Let P_S be the number of possible MT-EMs in $\boldsymbol{X}_S$. The upper bound of P_S is given by $P_1 \times P_2$. In real applications, it is unlikely that all possible transitions among land-cover classes in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ occur in an image pair. This is because changes usually occur with a small probability and because some transitions are unlikely or impossible from the physical point of view. However, P_S might be large enough to make it difficult to correctly identify all of them directly from the entire $\boldsymbol{X}_S$. This is especially true for MT-EM associated to changes that usually show a small prior probability with respect to the ones associated to no-change. Moreover endmembers in HS images may show a high variability [33], [34] due to the land-cover spectral properties and variable external factors (*e.g.*, atmospheric conditions, illumination, seasonal effects). Spectral signatures that belong to the same material transition may be different in different portions of the scene [34]. In this paper, we use a patch-scheme based on the local endmembers strategy [12], [35]. The patch scheme simultaneously handles both a possible large number of endmembers issue and small local spectral variability effects. Thus $\boldsymbol{X}_S$ is divided in Z regularly shaped patches (see Fig.3). Let $\boldsymbol{X}_{S,z}$ be the z-th (z=1,…,Z) patch of $\boldsymbol{X}_S$, where Z is the defined number of patches. Note that the parameter 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.

B. 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 $P_{S,z}$ ($\leq P_S$) of the MT-EMs; 2) The identified distinct MT-EMs $\boldsymbol{E}_{S,z}$ according to $P_{S,z}$.

The number $P_{S,z}$ of MT-EMs 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 [36], the noise whitened version of HFC (NWHFC) [37], the Hyperspectral Signal

Identification by Minimum Error (HySIME) method [38], and the recently proposed Eigenvalue Likelihood Maximization (ELM) algorithm [39]. Any endmember number estimation and endmember extraction algorithm can be used. 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 [39], which is totally parameter free and easy to be implemented. Then the popular Vertex Component Analysis (VCA) method [40] 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 [40].

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_U be the total number of MT-EMs in $\boldsymbol{U}$, $P_U = \sum_{z=1}^{Z} P_{S,z}$. Thus $\boldsymbol{U}$ is the set of all distinct MT-EMs extracted at the patch local scale. $\boldsymbol{U}$ is likely to include slightly different endmembers for the same land-cover transition, while representing the endmembers variability. All P_U 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 [7]. LMM has been intensively investigated in the literature [6], [7], [8], [34]. However, as mentioned in Section II, the original LMM is built based on the single-date image, whereas here a 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 change or no-change MT-EMs, which is modeled as:

$$\boldsymbol{x}_s(i,j)=\sum_{p=1}^{P_U} a_p(i,j)\boldsymbol{e}_p(i,j)+\boldsymbol{n}(i,j) \tag{4}$$

where $\boldsymbol{e}_p$ is the spectral signature of the p-th (p=1,…, P_U) 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. $\boldsymbol{A}_U$ and $\boldsymbol{U}$ are the set of a_p and $\boldsymbol{e}_p$ (p=1,…, P_U) for all pixels in $\boldsymbol{X}_S$, respectively. More complex linear or non-linear mixture models can be employed without loss of validity.

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

$$\begin{gathered}\tilde{\boldsymbol{A}}_U=\arg\min_{\boldsymbol{A}_U}\left\|\boldsymbol{X}_S-\boldsymbol{U}\boldsymbol{A}_U\right\|^2\\ \text{subject to: } \boldsymbol{A}_U\geq 0\end{gathered} \tag{5}$$

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

C. Change analysis

$\boldsymbol{U}$ includes all local MT-EMs that either belongs 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 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 $P_{U,c}$ and $P_{U,n}$ be the number of endmembers in $\boldsymbol{U}_c$ and $\boldsymbol{U}_n$, respectively, and $P_U=P_{U,c}+P_{U,n}$. 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 standard two-step approach to change detection was applied [2], [5], [24]. First binary change detection is performed to separate changed and no-changed MT-EMs, and then multiple classes of change are distinguished among the changed ones. In the first step, the magnitude (*i.e.*, the Euclidean distance) of the spectral difference (see Fig.2, two components in the extracted MT-EMs associated to $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, respectively) is analyzed according to change vector analysis:

$$\rho_{\boldsymbol{e}_p} = \sqrt{\sum_{b=1}^{B} \left(e_{p,B+b} - e_{p,b} \right)^2} \tag{6}$$

where $e_{p,b}$ is the b-th stacked channel of the p-th ($p = 1,\ldots, P_U$) endmember in $\boldsymbol{U}$. B is the number of the spectral channels in $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. It is known from the literature [2], [24], [25], [41] that the no-change class is associated to low magnitude values whereas changes are associated to high magnitude values. Thus, a given $\boldsymbol{e}_p$ is classified either into $\boldsymbol{U}_c$ or $\boldsymbol{U}_n$ depending on its magnitude value. A threshold T_ρ can be applied to the magnitude to separate the two classes. The threshold T_ρ is automatically estimated based on the histogram of the magnitude variable ρ calculated on $\boldsymbol{X}_D$ as [2], [24], [25], [41]:

$$\rho = \sqrt{\sum_{b=1}^{B} \left(\boldsymbol{X}_{D,b} \right)^2} \tag{7}$$

where $\boldsymbol{X}_{D,b}$ is the b-th (b=1,...,B) component of $\boldsymbol{X}_D$. Several thresholding algorithms can be found in the literature. The reader can refer to [42] for a survey. Here we used the approach proposed in [25] that exploits the Expectation Maximization (EM) algorithm for an unsupervised estimation of class probability density functions and the Bayesian decision theory for computing the minimum error decision threshold T_ρ according to the estimated probability density functions.

The MT-EMs in $\boldsymbol{U}_c$ are extracted from different patches representing the spatial distribution of different change targets on the image (*i.e.*, more than one MT-EM in $\boldsymbol{U}$ may be associated to the same land-cover transition). Thus in the second step grouping is required to cluster the MT-EMs that belong to the same change class while preserving their spectral variability. The final MT-EM groups represent the K unique change classes at a global level. For the endmember grouping, we adopted an iterative scheme that was designed based on the Spectral Angle Mapper (SAM) [12]. Instead of using SAM, in this work we selected the spectral measurement proposed in [43], 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 [44]:

$$\mathrm{SID}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) = \sum_{b=1}^{2B} r_b \log(r_b / m_b) + \sum_{b=1}^{2B} m_b \log(m_b / r_b) \tag{8}$$

and the SAM is defined as:

$$\mathrm{SAM}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) = \cos^{-1}\left[\sum_{b=1}^{2B} e_{\alpha,b} e_{\beta,b} \Bigg/ \sqrt{\sum_{b=1}^{2B} (e_{\alpha,b})^2 \sum_{b=1}^{2B} (e_{\beta,b})^2} \right] \tag{9}$$

So the SID-SAM combined spectral measure ϑ is defined as [43]:

$$\vartheta(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) = \mathrm{SID}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) \cdot \sin\left[\mathrm{SAM}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) \right] \tag{10}$$

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. The threshold T_ϑ is a user defined parameter that controls the similarity between endmembers required for grouping. Given the expected high similarity between MT-EMs of the same land-cover transition, T_ϑ is expected to be small. Finally, the endmember grouping results in K unique land-cover transitions $\left\{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}\right\}$. By considering also the no-change class ω_n, in total we have $K' = K + 1$ classes in $\boldsymbol{X}_S$.

D. Abundance combination and final change-detection map generation

Based on the grouping result, for each pixel in $\boldsymbol{X}_S$, 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 $\boldsymbol{A}_{U,\omega_\varepsilon}$ of class ω_ε in $\boldsymbol{X}_S$ is computed as:

$$\boldsymbol{A}_{U,\omega_\varepsilon} = \sum_{\boldsymbol{e}_p \in \omega_\varepsilon} \boldsymbol{A}_{U,\boldsymbol{e}_p} \tag{11}$$

where $\boldsymbol{A}_{U,\boldsymbol{e}_p}$ is the abundance map of a given MT-EM $\boldsymbol{e}_p$ in $\boldsymbol{U}$, and $\boldsymbol{e}_p \in \omega_\varepsilon$, $p=1,\ldots, P_U$. Abundances in equation (11) solve the unmixing problem and provide land-cover transition information at sub-pixel level.

As a side output a crisp multiple-change detection map can be generated by assigning to each pixel $x_S(i, j)$ $(1\le i \le I, 1\le j \le J)$ in $\boldsymbol{X}_S$ the class ω_ε having the maximum abundance value:

$$x_s(i,j) \in \arg\max_{\omega_\varepsilon \in \Omega} \left(a_{\omega_\varepsilon}(i,j) \right) \tag{12}$$

where $a_{\omega_\varepsilon}(i,j)$ is the abundance value of class ω_ε in pixel $x_S(i,j)$. This output can be useful in the context of some applications or for comparison with the results of state-of-the-art methods working at pixel level.

IV. EXPERIMENTAL RESULTS

A. Simulated stacked multitemporal hyperspectral data set

The first data set is defined by simulating a stacked multitemporal HS data set using reference MT-EMs from USGS digital spectral library-splib06a [45]. Three spectral endmembers (*i.e.*, plywood_gds365, sheetmetal_gds352, vinyl_plastic_gds372) were selected and their dimensionality was resampled to 200 channels (see Fig.4). We stacked the known endmembers to simulate nine MT-EMs, *i.e.*, nine class transitions. Six are associated to changes and three to no changes. MT-EMs spectral signatures are provided in Fig.5 (dashed lines). The 9 MT-EMs were employed to generate a 300×300 pixels image. Each pixel is associated to a 400-dimensional feature vector computed as linear mixture of the nine MT-EMs. Four versions of the stacked image were generated with different levels of Signal-to-Noise Ratio (SNR) values (*i.e.*, SNR=10, 20, 30, 40 dB).

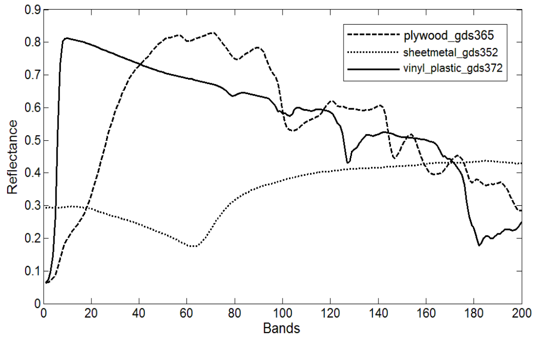


Fig.4 Spectral endmembers from USGS spectral library [45] used for the generation of the simulated data sets.

The performance of the proposed MSU approach has been quantitatively evaluated by computing the Root Mean Square Error (RMSE) [46] value between the class abundance estimated by the proposed MSU method and the known simulated abundance values:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(a_{\omega_\varepsilon}(i,j) - a_{Ref,\omega_\varepsilon}(i,j)\right)^2}{IJ}} \tag{13}$$

where $a_{Ref,\omega_\varepsilon}(i,j)$ is the abundance value in the spatial position (*i*,*j*) in the simulated abundance map $\boldsymbol{A}_{ref,\omega_\varepsilon}$ of class ω_ε. The average RMSE value among all classes was used for comparison. For each noise level, several trials were conducted by: i) varying T_ϑ in the range [0.005, 0.050] with a step of 0.001; and ii) considering four values for the *Z* parameter, *i.e.*, 1 (the entire image), 2, 4 and 8. All experiments were carried out by using Matlab R2014a, on an Intel(R) Core(TM) i5-4200 quad-core 2.50GHz PC with 4GB of RAM. Time consumption is provided in Table I.

For each simulated data set, a sensitive analysis was conducted with respect to the value of T_ϑ that controls the grouping of MT-EMs. Since T_ϑ optimal range is independent on *Z*, we report for each data set the results for *Z*=8. As T_ϑ increases the number of MT-EMs decreases. By increasing the signal-to-noise ratio the method identifies the correct *K*=6 number of changes when T_ϑ is in the ranges [0.017, 0.020], [0.013, 0.017], [0.011, 0.015] and [0.012, 0.014], respectively. Thus the optimal T_{ϑ} is in the range [0.01, 0.02], approximately. This range can be used as a guidance for an empirical selection of the optimal threshold. However, T_{ϑ} depends on the considered data set and the complexity of the scenario. For each SNR level we selected a threshold value in the above range (*i.e.*, 0.018, 0.015, 0.013 and 0.013) for the next experiments.

Table I summarizes MSU performance when varying Z. We can observe that: 1) by increasing the parameter Z (*i.e.*, number of patches), more local MT-EMs were identified (see the increasing number of P_U in all data sets). Thus the correct number of change classes (*i.e.*, K=6) are more likely to be detected. MSU always fails to detect the correct K value when the entire image is considered (*i.e.*, Z=1) thus confirming the effectiveness of the patch scheme; 2) by increasing the amount of noise, also the Z required to detect the correct K value increases. The higher number of patches allows to better capture the local effects of noise. For example, in the case of SNR=10dB, Z should be set to 8 to detect all changes, whereas in the case of SNR=40dB, Z=2 is enough; 3) as expected the average multi-class RMSE value increases when increasing the noise level. In other words it becomes more difficult to properly estimate MT-EMs. However, in all cases the RMSE values are low, thus a good approximation of the abundances is still achieved. The average RMSE performance saturates when Z is large enough to detect the correct K value. However the computational time increases significantly when increasing Z; 4) the time cost of the proposed MSU is in general small.

In addition, a quantitative evaluation is conducted by comparing the estimated MT-EMs with the corresponding reference ones. Fig.5 shows an example of the estimated MT-EMs for the SNR=30dB ($Z=4$ and $T_{\vartheta}=0.013$) experiment. The estimated MT-EMs represent six change classes having unique spectral shapes. After the grouping step, each group of MT-EMs shows a high similarity to one of the simulated MT-EMs, belonging either to a change or a no-change class (see Fig.5).

TABLE I CD RESULTS OBTAINED BY THE PROPOSED MSU METHOD. THE AVERAGE RMSE AND THE COMPUTATIONAL TIME ARE GIVEN FOR THE CASES IN WHICH THE NUMBER OF CHANGES IS CORRECTLY IDENTIFIED (SIMULATED STACKED MULTITEMPORAL HYPERSPECTRAL DATA SET).

SNR (dB)	***Z***	***P_U***	***K***	**Average Multi-class RMSE**	**Computational Time (s)**
10	1	6	4	-	-
	2	11	4	-	-
	4	26	5	-	-
	8	46	**6**	6.093	113.99
20	1	6	4	-	-
	2	12	4	-	-
	4	27	**6**	**5.178**	78.66
	8	48	**6**	5.466	126.74
30	1	6	4	-	-
	2	12	5	-	-
	4	29	**6**	**4.425**	96.16
	8	53	**6**	4.473	121.47
40	1	6	4	-	-
	2	12	**6**	4.395	59.37
	4	29	**6**	**4.302**	81.62
	8	61	**6**	4.308	144.91

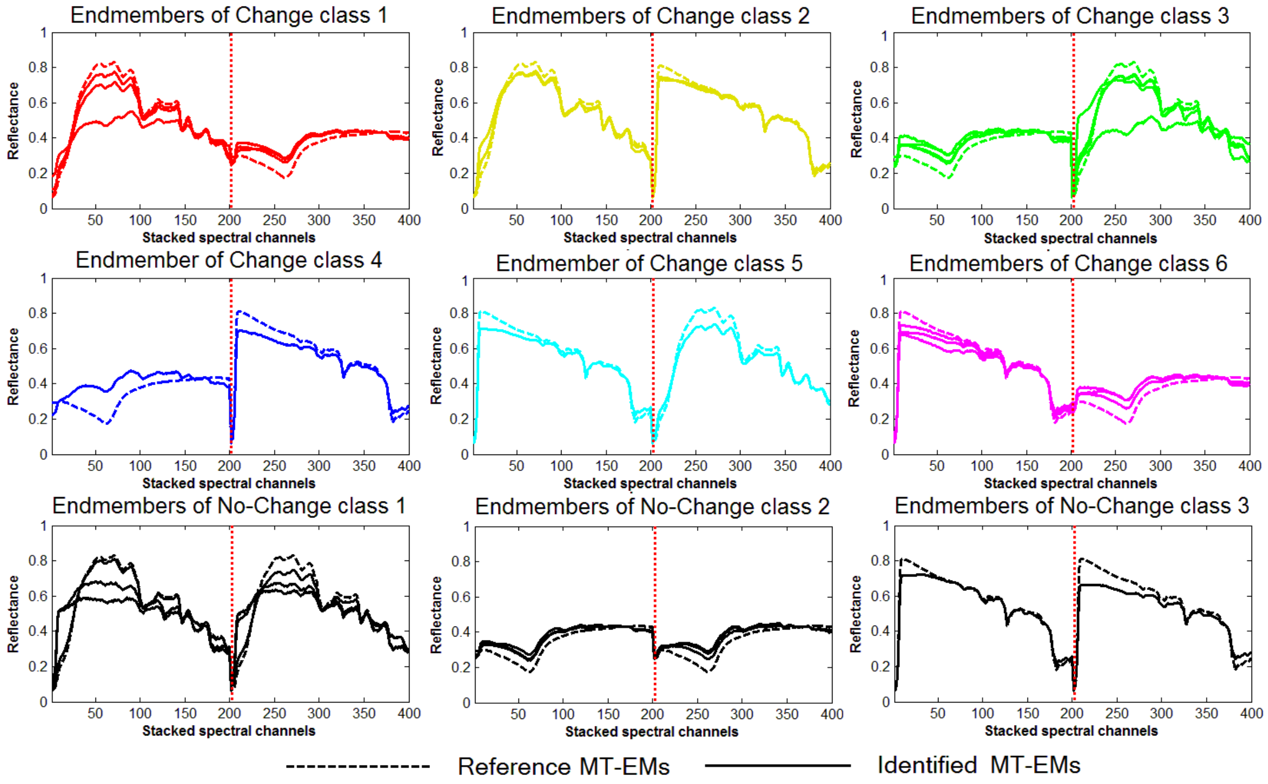
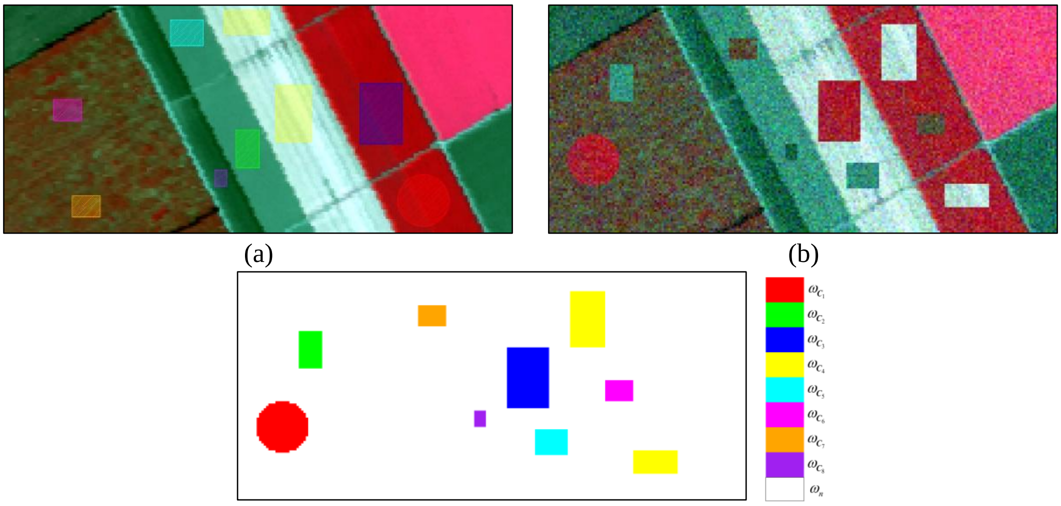


Fig.5 Spectral signatures of the MT-EMs (solid line) extracted by the proposed MSU approach (Z=4, T_g=0.013) on the simulated data set (SNR=30dB), where six unique change classes (in different colors) are identified, and three no-change classes are in black. The reference MT-EMs are plotted as dashed lines.


(a) (b)

(c)

Fig.6 False color composites (bands: R: 40, G: 30, B: 20) of (a) the HS image acquired by the AVIRIS sensor in Salinas scenario (X_1) and (b) the simulated changed image (X_2) computed with an additive white Gaussian noise (SNR=20dB). (c) pixel-level change reference map (eight changes in different colors, and no-change class in white color).

B. Simulated hyperspectral remote sensing data set

The second data set is made up of a HS remote sensing 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). A subset of the whole image was selected having a size of 217×97 pixels. In pre-processing, 20 water absorption bands (*i.e.*, bands 108-112, 154-167 and 224) were discarded, obtaining 204 bands for the 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 (a)) and inserted back in different spatial positions on $\boldsymbol{X}_1$ by replacing the whole spectral vectors. Thus $\boldsymbol{X}_2$ was generated with eight simulated change classes (see Fig.6 (b)). A small constant bias value was applied to $\boldsymbol{X}_2$ to simulate a stationary radiometric difference. White Gaussian noise was added with different levels of SNR values (*i.e.*, SNR=10, 20, 30, 40dB). 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 (a) and (b), respectively. Fig.6 (c) is the reference change map. Detailed simulated land-cover transitions and their corresponding number of samples are listed in Table II.

TABLE II

SIMULATED CHANGE CLASSES AND RELATED NUMBER OF SAMPLES (SALINAS DATA SET)

Change class	**Simulated changes (from X_1 to 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

MSU was applied to the four image pairs leading to similar results. Let us analyze in detail the ones obtained for the image pair associated with SNR=20dB. Image $\boldsymbol{X}_S$ was divided it into Z = 4 patches $\boldsymbol{X}_{S,z}$ (z = 1,…, 4) (Z = 1 and Z = 2 failed to detect all the simulated changes), see Fig.7. $\boldsymbol{U}$ was computed as the union of all MT-EMs from all patches (see Fig.7) and their abundances $\boldsymbol{A}_S$ were calculated according to (5). $\boldsymbol{U}_c$ and $\boldsymbol{U}_n$ were separated in $\boldsymbol{U}$ by automatic thresholding. The estimated T_ρ was obtained by [25] and resulted equal to 0.82. According to the analysis presented in Section IV.A T_ϑ was set to 0.015 and resulted as expected in K=8 changes. MT-EM abundances were summed into the final abundances of change and no-change classes based on the grouping results to generate the final subpixel CD maps (11) and the crisp pixel-level CD map (12).

From Fig.7 we can see that the MSU approach extracted 35 MT-EMs from four patches, representing very well the distinct endmembers (including both the change and no-change classes) in each local patch. 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 correct eight unique change classes. The spectral signatures of the detected MT-EMs (associated with the detected eight change classes and the no-change class) are shown in Fig.8, and the corresponding abundances of each class are illustrated in Fig.9 (a)-(i).

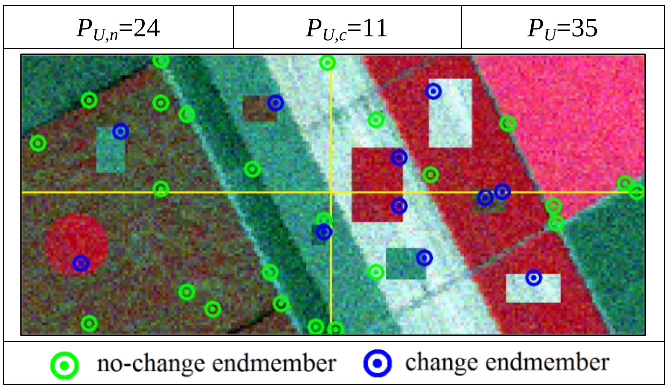


Fig.7 MSU estimated MT-EMs in the simulated data set with SNR=20dB (Z=4, T_g=0.015). From left to right, up to down the four patches $\boldsymbol{X}_{S,1}$ ($P_{S,1}$=10), $\boldsymbol{X}_{S,2}$ ($P_{S,2}$=8), $\boldsymbol{X}_{S,3}$ ($P_{S,3}$=9) and $\boldsymbol{X}_{S,4}$ ($P_{S,4}$=8) are shown.

From Fig.8 and Fig.9 we can observe that: 1) the detected eight change classes have unique spectral signatures in the $\boldsymbol{X}_S$ domain. 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 both components; 2) MT-EMs are identified from spatial different regions in $\boldsymbol{X}_S$, but group 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 to properly describe the unique change classes in the global scene; 3) the abundance maps show a good unmixing and separation result among classes (see Fig.9 (a)-(i)), where the detected eight changes and the no-change background shows a clear contrast of their abundances (with respect to the colors). Thus the considered multiple change-detection problem was successfully solved by estimating the percentage of class substances at subpixel level. Similar considerations hold for all the four simulated pairs.

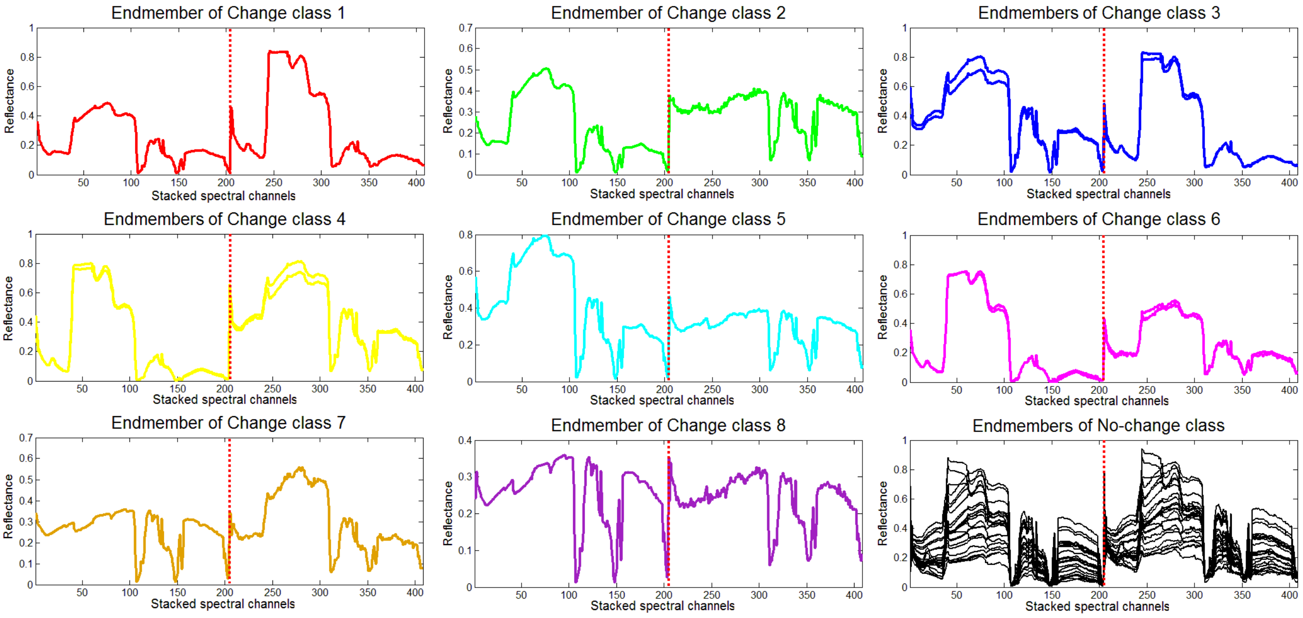


Fig.8 Spectral signatures of the MT-EMs extracted by the proposed MSU approach on the simulated data set (SNR=20 dB, Z=4, T_g=0.015), where eight unique change classes (in different colors) are identified, and the no-change class endmembers are in black. Endmembers that belong to the same class represent the variability in the image.

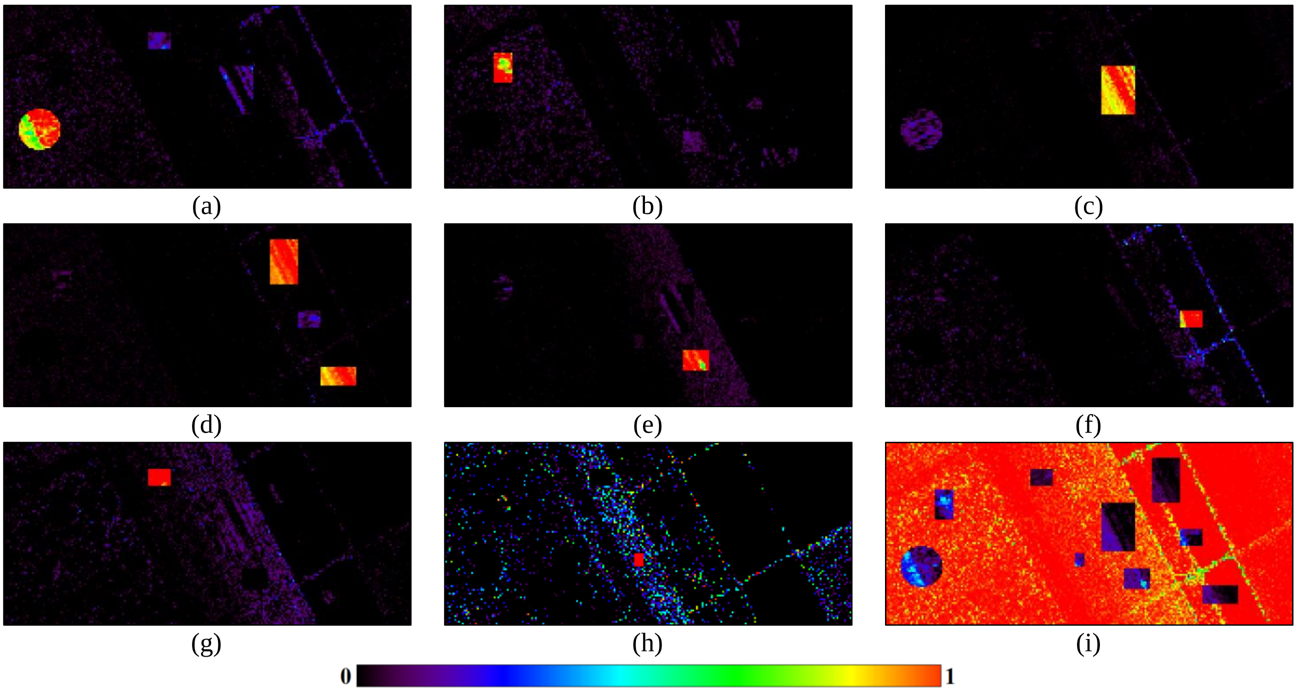

(a) (b) (c)
(d) (e) (f)
(g) (h) (i)

Fig.9 Final class abundances on the simulated data set (SNR=20dB, Z=4, T_g=0.015), where (a)-(h) are the abundances of eight change classes, and (i) of the no-change class.

Pixel-level CD maps obtained by the proposed MSU (with Z = 4) on four simulated HS data sets are shown in Fig.10. From a qualitative analysis of the CD maps, we can observe that the proposed MSU achieved good 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.10, 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 does not guarantee a good local analysis. In this case, a smaller patch scale is required to reach the optimal detection scale.

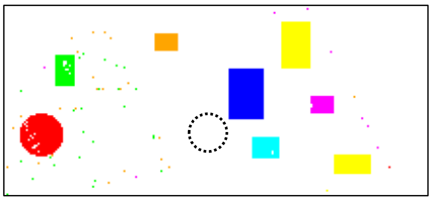
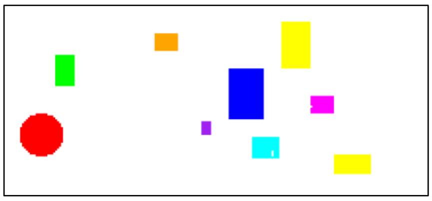
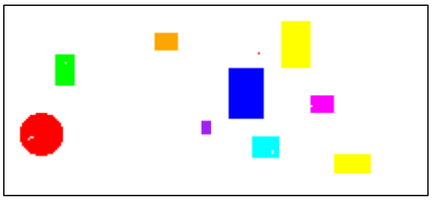
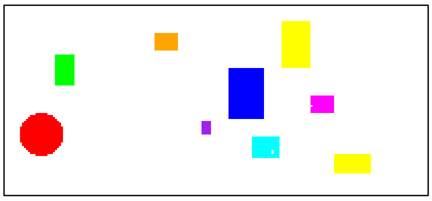
SNR (dB)	CD maps	SNR (dB)	CD maps	
10		30		🟥 ω_{C_1} 🟩 ω_{C_2} 🟦 ω_{C_3} 🟨 ω_{C_4} ■ ω_{C_5} ■ ω_{C_6} 🟧 ω_{C_7} 🟪 ω_{C_8} ⬜ ω_n
20		40		

Fig.10 Change-detection maps obtained by the proposed MSU approach on different simulated HS 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.

As no sub-pixel reference information is available for this data set, quantitative analysis is conducted at a pixel level. The pixel-level CD maps obtained by the proposed MSU approach were compared with those yielded by three pixel-level CD-HS methods: 1) the hierarchical spectral change vector analysis (HSCVA) [2], 2) the sequential spectral change vector analysis (S^2CVA) [5], and 3) the unsupervised *k*-means clustering applied to the changed spectral change vectors. According to Sec. IV.A, T_ϑ was set to 0.015 for all four cases. Four Z values were tested (*i.e.*, Z=1, 2, 4, 8) to detect the smallest one that guarantees the detection of the eight simulated changes. The optimal Z equals to 4 in the cases when SNR=20, 30, 40dB, and Z=8 in the case of SNR=10dB. Advantage was given to the *k*-means by providing 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 as the average of 100 trails in order to reduce the uncertainty of the random initialization.

TABLE III

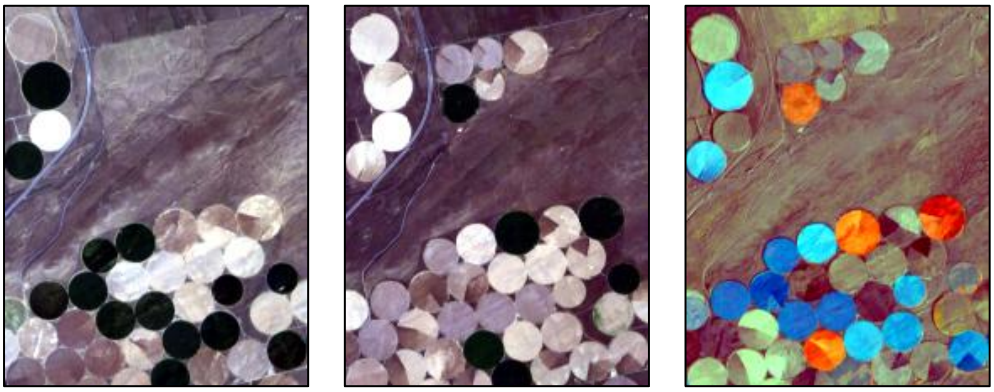
CHANGE DETECTION ACCURACY, ERROR INDICES AND TIME COST OBTAINED BY THE CONSIDERED METHODS (DIFFERENT SIMULATED HS DATA SETS).

SNR	**Method**	**Unsupervised and automatic detection?**	***OA* (%)**	***Kappa***	**Errors (pixel)**	**Time Cost (sec.)**
10 dB	HSCVA	No	99.69	0.9829	64	312.67
	S^2CVA	No	99.74	0.9856	54	196.18
	k-means	No	98.32	0.8988	353	47.93
	MSU (Z=8, T_ϑ=0.015)	Yes	99.54	0.9749	96	71.42
20 dB	HSCVA	No	99.99	0.9995	2	304.66
	S^2CVA	No	99.98	0.9989	4	175.41
	k-means	No	99.47	0.9702	111	49.33
	MSU (Z=4, T_ϑ=0.015)	Yes	99.96	0.9978	8	62.69
30 dB	HSCVA	No	99.99	0.9997	1	297.54
	S^2CVA	No	99.98	0.9989	4	177.59
	k-means	No	99.66	0.9811	71	54.61
	MSU (Z=4, T_ϑ=0.015)	Yes	99.98	0.9989	4	65.77
40 dB	HSCVA	No	99.99	0.9997	1	316.76
	S^2CVA	No	99.98	0.9989	4	182.25
	k-means	No	99.78	0.9880	45	53.48
	MSU (Z=4, T_ϑ=0.015)	Yes	99.99	0.9992	3	66.81

Quantitative results are given in Table III, 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. From Table III, one can see that the two state-of-the-art hierarchical methods achieved good results in all four cases, and 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 [2], [5]. 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 [2] and a user interaction for change identification in S^2CVA [5] are required, respectively); 2) A significant effort is required to search for the hierarchical structure, which increases the implementation complexity and the time cost (see Table III). For example, both hierarchical methods resulted in a three-level hierarchy with more than twelve nodes and a time consumption of 3-5 minutes for the simulated data set associated with SNR=20dB. On the contrary, the proposed automatic MSU method completed the process in around 1 minute with an accuracy decrement of only 0.02%. Note that although the reference *k*-means algorithm was applied using the known number of classes and relatively computational fast, it resulted in a significantly higher number of errors than the other three methods (see Table III). This indicates the difficulty of the CD-HS problem to be solved when applying the clustering directly to the high-dimensional HS images.

C. Real Hyperion hyperspectral remote sensing data set

This data set is made up of a pair of real bi-temporal HS 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.*, 8-57, 82-119, 131-164, 182-184, 187-220) out of the original 242 bands were used for the considered CD task. 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.11 (a) and (b) show the false color composite of the $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ images, respectively. Fig.11 (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.11 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 (X_1), (b) 2007 (X_2), and (c) composite of three X_D channels (wavelength: R: 823.65nm, G: 721.90nm, B: 620.15nm).

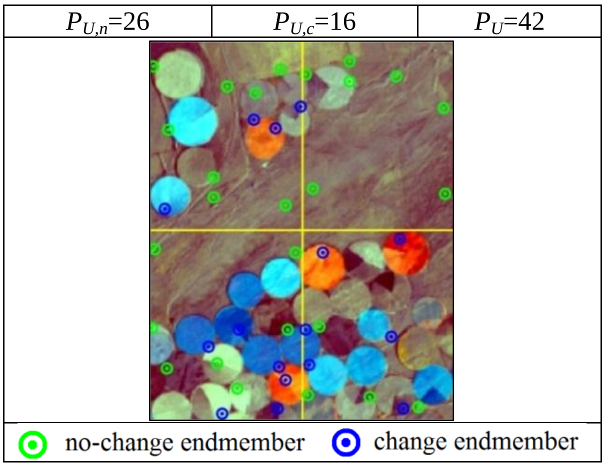


Fig.12 MT-EMs extracted by the proposed MSU approach (Z=4, T_ϑ =0.013) on the real bitemporal Hyperion data set. From left to right, up to down the four patches $X_{S,1}$ ($P_{S,1}$=12), $X_{S,2}$ ($P_{S,2}$=7), $X_{S,3}$ ($P_{S,3}$=13) and $X_{S,4}$ ($P_{S,4}$=10) are shown.

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.12). MSU identified 42 local MT-EMs (Fig.12). T_ρ was automatically detected and equals 1.486. T_ϑ was set according to the analysis presented in Section IV.A equal to 0.013. Seven unique change classes were detected. The final class abundances and CD map were computed according to (11) and (12).

The extracted MT-EMs by MSU are illustrated in Fig.13. Their corresponding abundances are shown in Fig.14 (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 $\boldsymbol{X}_S$, respectively. The MT-EMs of the change classes have discriminable and unique spectral signatures among each other (See Fig.13 change classes 1-7); 2) an endmember set rather than a single endmember was used to represent a detected change class, thus the change targets were described via unmixing by also considering the endmember variability (see Fig.13 change classes 1-3, 5-7); 3) even though we are not interested in distinguishing MT-EMs of no-change classes among each other $\boldsymbol{U}_n$, is an important and non-negligible source of information for the mixture model in $\boldsymbol{X}_S$. They were considered in the unmixing process, and their abundances summed into the final no-change class ω_n. A good estimation of the no-change background indicates its good separation from the change classes (see Fig.14 (h)); 4) the abundances of change classes (see Fig.14 (a)-(g)) confirm the accurate representation 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. The 7th MT-EM in Fig.14 (g) is characterized by a set of circular patterns. They are likely to be generated by some residual mis-registration. This induces the detection of false alarms along borders between roads surrounding the agricultural fields and the fields their self. Visual analysis on the Google maps [47] tells us that those roads have an average width of 6-10m, which is less than half pixel on the Hyperion images whose spatial resolution is 30m. Despite in the preprocessing step the residual error of co-registration is limited within 0.5 pixel, it 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 [48].

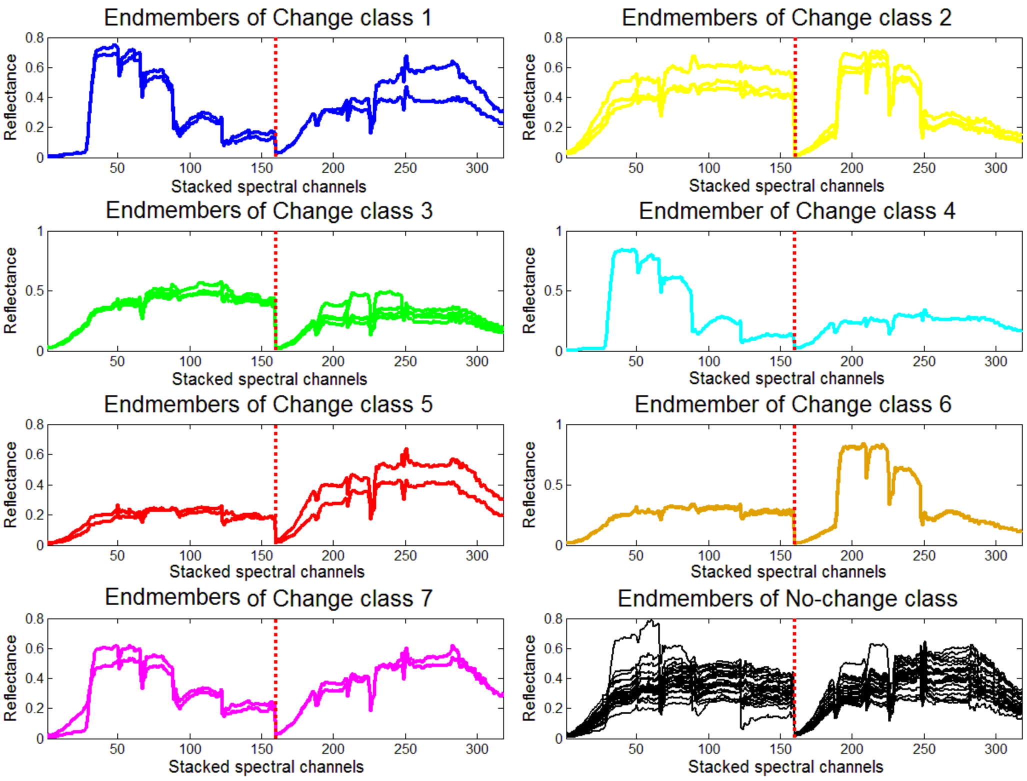
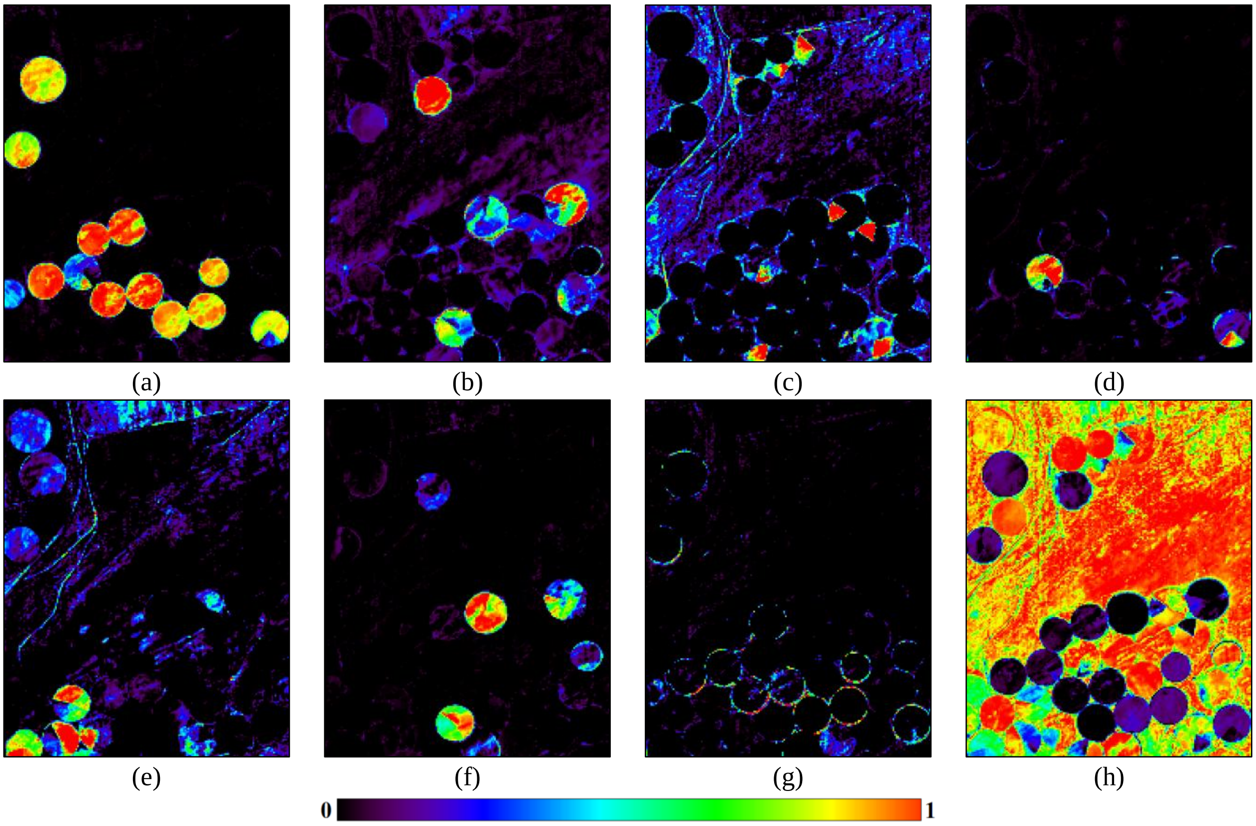


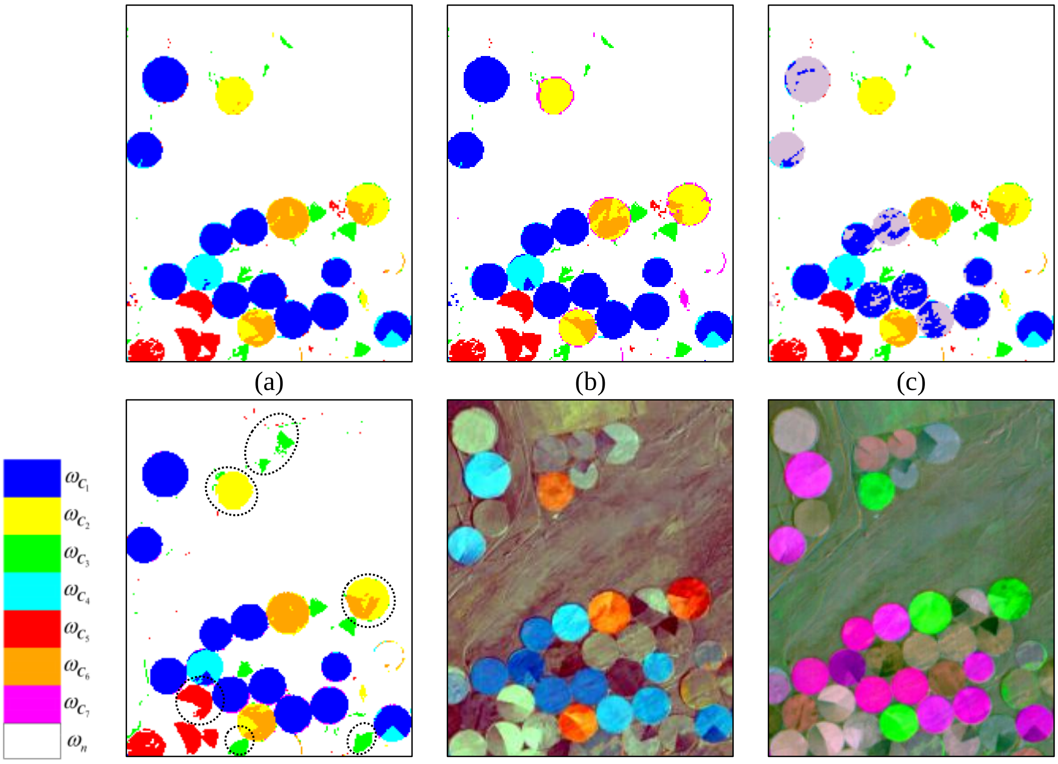
Fig.13 Spectral signatures of MT-EMs extracted by the proposed MSU approach (Z=4, T_ϑ =0.013) on the real bi-temporal Hyperion HS data set. Seven unique change classes are identified and shown in different colors, while the no-change MT-EMs are in black. Endmembers that belong to the same class represent the variability in the image.



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

(e) (f) (g) (h)

Fig.14 Final class abundances obtained by the proposed MSU approach (Z=4, T_ϑ =0.013) on the real bi-temporal Hyperion data set: (a)-(g) show the seven change classes and (h) the no-change class.



(a) (b) (c)

(d) (e) (f)

Fig.15 Change-detection maps obtained by: (a) HSCVA [2] (K=6); (b) S^{2}CVA [5] (K=7); (c) *k*-means (K=7) and (d) 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 X_D are provided for visual comparison in: (de) wavelength: R: 823.65nm, G: 721.90nm, B: 620.15nm; (f) R: 1729.7nm, G: 752.43nm, B: 548.92nm.

CD maps obtained by the considered approaches are shown in Fig.15 (a)-(d). The two reference hierarchical methods detected six and seven kinds of changes, respectively. The proposed MSU (Z=4, T_{ϑ}=0.013) detected K=7 change classes like S^{2}CVA. It can be clearly seen that change targets were better recognized in the result of MSU (Fig.15 (d)) than in those of the two hierarchical methods (Fig.15 (a) and (b)) and the *k*-means (see Fig.15 (c)). Both borders and homogenous areas are better detected. Comparing to the reference false color composite of $\boldsymbol{X}_D$ bands in Fig.15 (e) and (f), one can observe that the changes highlighted in Fig.15 (d) (especially the change class ω_{C_3} in green color) were detected more accurately. Accordingly, the multitemporal spectral mixture model used by MSU better captures the nature of changes by a detailed investigation of the spectral signature in $\boldsymbol{X}_S$, whereas standard crisp reference methods tend to under estimate changes. 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.15 (c)). It is important to note that the proposed MSU was implemented in an automatic way. However, it detected successfully the multiple change classes and resulted in slightly better results than the other techniques without any end-user participation.

V. DISCUSSION AND CONCLUSION

In this paper, a novel multitemporal spectral unmixing (MSU) approach has been proposed to address the challenging multiple change detection problem in multitemporal HS 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 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 HS 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 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 to a given class with respect to a crisp decision based on the pure-pixel theory. Accordingly, more sub-pixel 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 MSU method is automatic and unsupervised, thus allowing a fast and efficient CD without requiring ground truth training data. Two parameters should be manually selected, namely Z and T_{ϑ}. However, as described in this paper their selection is relatively an easy task. Once fixed, MSU leads to better performance than the state-of-the-art CD-HS techniques and guarantees a higher efficiency. Note that other endmember extraction techniques or unmixing models can be adopted within the proposed MSU approach.

3) Differently from the state-of-the-art multitemporal unmixing methods mainly based on post-unmixing comparison, the proposed approach explicitly considers temporal correlation and does not require training samples.

Future developments will include: 1) study an automatic technique for a robust search for the optimal parameters in MSU; 2) investigation of the advanced models (*e.g.*, sparse model) for improving the unmixing result; and 3) development of other unsupervised multitemporal unmixing approaches (*e.g.*, strategies based on unsupervised post-unmixing comparison). It is worth noting that in the latter case many challenges should be considered like the fact that: the number of endmembers on two images may be different, leading to the difficulties in endmember determination or matching; and the same endmember may show different spectra in two times images.

REFERENCE

[1] 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.

[2] 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.

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

[4] 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.

[5] S. Liu, L. Bruzzone, F. Bovolo, M. Zanetti, and P. Du, "Sequential Spectral Change Vector Analysis for Iteratively Discovering and Detecting Multiple Changes in Hyperspectral Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 8, pp. 4363-4378, 2015.

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

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] D. Lu, M. Batistella, and E. Moran, "Multitemporal spectral mixture analysis for Amazonian land-cover change detection," *Canadian Journal of Remote Sensing*, vol. 30, no. 1, pp. 87-100, 2004.

[15] 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.

[16] J. Chen, X. Chen, X. Cui, and J. Chen, "Change Vector Analysis in Posterior Probability Space: A New Method for Land Cover Change Detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 2, pp. 317-321, Mar. 2011.

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

[19] 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.

[20] A. P. Tewkesbury, A. J. Comber, N. J. Tate, A. Lamb, and P. Fisher, "A critical synthesis of remotely sensed optical image change detection techniques," *Remote Sensing of Environment*, vol. 160, p. 1–14, Apr. 2015.

[21] P. J. Howarth and G. M. Wickware, "Procedure for change detection using Landsat," *International Journal of Remote Sensing*, vol. 2, pp. 277-291, 1981.

[22] L. Yang, G. Xian, M. J. Klaver, and D. Brian, "Urban Land-Cover Change Detection through Sub-Pixel Imperviousness Mapping Using Remotely Sensed Data," *Photogrammetric Engineering & Remote Sensing* , vol. 69, no. 9, pp. 1003-1010, 2003.

[23] G. M. Foody and H. T. X. Doan., "Variability in soft classification prediction and its implications for sub-pixel scale change detection and super resolution mapping," *Photogrammetric Engineering & Remote Sensing*, vol. 73, no. 8, pp. 923-933, 2007.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] L. Bruzzone, D. F. Prieto, and S. B. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 37, no. 3, pp. 1350-1359, 1999.

[33] 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.

[34] 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.

[35] L. Drumetz, et al., "Binary partition three-based local spectral unmixing," in *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS 2014)*, 2014, pp. 1-4.

[36] 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.

[37] 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.

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

[39] 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.

[40] 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.

[41] 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.

[42] 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.

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

[44] 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.

[45] R. N. Clark, et al. (2007) USGS digital spectral library splib06a: U.S. Geological Survey, Digital Data Series 231. .

[46] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315--1326, 1999.

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

[48] 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.