# Improving Change Detection in Forest Areas Based on Stereo Panchromatic Imagery Using Kernel MNF

Jiaojiao Tian, Allan Aasbjerg Nielsen, and Peter Reinartz, *Member, IEEE*

***Abstract*—The goal of this paper is to develop an efficient method for forest change detection using multitemporal stereo panchromatic imagery. Due to the lack of spectral information, it is difficult to extract reliable features for forest change monitoring. Moreover, the forest changes often occur together with other unrelated phenomena, e.g., seasonal changes of land covers such as grass and crops. Therefore, we propose an approach that exploits kernel Minimum Noise Fraction (kMNF) to transform simple change features into high-dimensional feature space. Digital surface models (DSMs) generated from stereo imagery are used to provide information on height difference, which is additionally used to separate forest changes from other land-cover changes. With very few training samples, a change mask is generated with iterated canonical discriminant analysis (ICDA). Two examples are presented to illustrate the approach and demonstrate its efficiency. It is shown that with the same amount of training samples, the proposed method can obtain more accurate change masks compared with algorithms based on k-means, one-class support vector machine, and random forests.**

***Index Terms*—Change detection, digital surface model (DSM), forest, kernel Minimum Noise Fraction (kMNF), optical stereo data.**

## I. Introduction

FOREST management and observation are important and time-consuming tasks. Automatic inventory and monitoring of forest changes have drawn the interest of many agencies, particularly after wind breakage. Satellite data are a valuable data source from which change information can be efficiently extracted for large regions. However, automatically extracting changes from satellite images is not easy, particularly when forest changes are mixed with other changes.

Forest change detection studies in remote sensing mainly involve the use of multispectral data. Existing change detection techniques have been well summarized in [1]–[3]. In the early stages, vegetation index differencing is one of the most important methods for forest change detection, as the vegetation covers can be well highlighted through band-by-band arithmetic calculation [1], [4], [5]. In order to fully use change features from all channels, change vector analyses (CVAs) [6]–[8] were proposed to analyze more change features. Principal component analysis (PCA) [1], [9] and multivariate alteration detection (MAD) [10] were used to highlight changes by transforming multi- or hyperspectral data in other feature space. In recent years, more techniques such as IR-MAD [11] and kernel-based change detection methods [12]–[16] have been reported to improve change detection accuracy and efficiency. Good general references for kernel methods are [17], [18], the latter with a remote sensing focus.

However, if only panchromatic images are available, the aforementioned methods cannot be used, and forest change detection becomes more challenging. For instance, the Cartosat-1 satellite (IRS-P5), which was launched by the Indian Space Research Organization in 2005 [19], [20], can only acquire panchromatic images but, at the same time, delivers stereo image data to derive digital surface models (DSMs). Moreover, in other cases, only panchromatic data are available, due to sensor or budget restrictions, and therefore, it is of interest to develop a methodology using only these data still able to produce reliable forest change masks. The limitation of such panchromatic-image-based change detection is mainly due to the low spectral differences between object classes. Different classes such as tree species and grassland often overlap in their gray-value distributions. In particular, forest regions are, in most cases, surrounded by crops or meadows, which are subject to more frequent changes and are always sensitive to seasonal changes [19]. Thus, the height can be very helpful in separating forest changes (particularly clear cuts, wind breakage, and growth) from other unrelated changes.

In this paper, a new change detection methodology is proposed. It combines the height changes from stereo images and gray-value changes from panchromatic images. Then, kernel Minimum Noise Fraction (kMNF) [13] is adopted to transform these two features into high-dimensional feature space via kernel functions; thus, MNF can be used to highlight the forest changes. Once the change map is generated, iterated canonical discriminant analysis (ICDA) [21] is applied to produce a forest change mask. In addition, this paper is aiming at studying the effectiveness of the proposed approach based on Cartosat-1 data and to compare its performance to other commonly used change classification techniques based on k-means, one-class support vector machine (OSVM), and random forests.

## II. Method

### A. Height Extraction

When stereo imagery is available, a DSM can be generated [19], [22], and the contained height information can be used

Manuscript received June 4, 2013; revised September 26, 2013 and January 14, 2014; accepted February 6, 2014. Date of publication April 16, 2014.

J. Tian and P. Reinartz are with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen 82234, Germany (e-mail: jiaojiao.tian@dlr.de; peter.reinartz@dlr.de).

A. A. Nielsen is with the Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark, 2800 (e-mail: alan@dtu.dk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2014.2308012

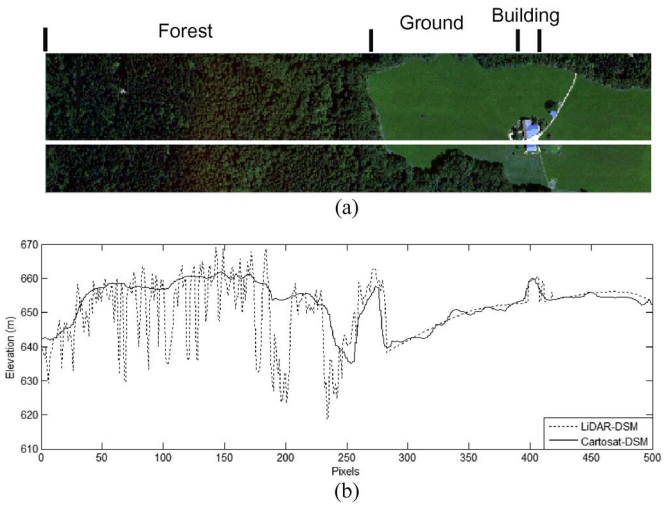


Fig. 1. Profile comparison of a DSM from Cartosat-1 and a LiDAR-DSM.

as an additional change feature for forest change detection. The quality of the generated DSMs is mainly constrained by the stereo matching technique. There are many matching algorithms available, but the dense matching methods are the most powerful in generating high-quality DSMs [22]. In this paper, semi-global matching (SGM) [19], [22] is adopted due to its good performance and easy implementation. Details on the rational polynomial coefficient correction for georeferencing and DSM refinement is illustrated in our earlier research work [19]. After obtaining DSMs from the two dates, further coregistration between them is necessary to remove any shift in three dimensions that might exist [19].

Fig. 1 presents the quality of the generated DSM in forest regions by comparing it with a DSM from LiDAR scanning data (LiDAR-DSM). The dashed line and the solid line in Fig. 1(b) represent the profile of the LiDAR-DSM and the Cartosat-DSM, respectively, along the white line in Fig. 1(a). For this example, we analyze an area in Bavaria that is different from that used in the experimental part, as the availability of LiDAR data is limited to this data set. It can be seen that, although the Cartosat-DSM is not able to get a surface height in between trees as detailed in the LiDAR-DSM, the extracted canopy height matches well the LiDAR-DSM in the forest region. Therefore, changes in canopy height as deforestation can be accurately highlighted by subtracting two Cartosat-DSMs. The main problems of DSM subtraction results are the blurred boundaries of trees and false alarms deriving from randomly distributed incorrect height information. More details on DSM quality evaluation can be found in [23]. In the following, we will focus on highlighting forest changes based on DSMs and panchromatic image data.

### *B. kMNF Analysis*

In (linear) minimum noise fraction (MNF) analysis [24], we consider multivariate measurements represented by the $n$ by $p$ data matrix $\boldsymbol{X}$ (one row per observation or pixel, one column per variable or spectral band) as being the sum of an uncorrelated signal $\boldsymbol{X}_S$ and noise $\boldsymbol{X}_N$, $\boldsymbol{X} = \boldsymbol{X}_S + \boldsymbol{X}_N$. With uncorrelated signal and noise, the variance–covariance matrix of $\boldsymbol{X}$, $\boldsymbol{S}_X$ is equal to the sum of the variance–covariance matrices of the signal $\boldsymbol{S}_S$ and the noise $\boldsymbol{S}_N$, $\boldsymbol{S}_X = \boldsymbol{S}_S + \boldsymbol{S}_N$. We then project the originally measured variables $\boldsymbol{X}$ onto the directions $\boldsymbol{a}$ in the feature space, which minimize the noise fraction $NF$ (or maximize $1/NF$). $NF$ is defined as the ratio between the variance of the projected noise and the variance of the projected total

$$NF = \frac{\boldsymbol{a}'\boldsymbol{S}_N\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{S}_S\boldsymbol{a}} = \frac{\boldsymbol{a}'\boldsymbol{X}'_N\boldsymbol{X}_N\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{a}} \tag{1}$$

where $\boldsymbol{X}$ and $\boldsymbol{X}_N$ are column centered. A regularized version of $1/NF$ is

$$\frac{1}{NF} = \frac{\boldsymbol{a}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{a}}{\boldsymbol{a}'\left[(1-\lambda)\boldsymbol{X}'_N\boldsymbol{X}_N + \lambda\boldsymbol{I}\right]\boldsymbol{a}} \tag{2}$$

where $\boldsymbol{I}$ is the unit matrix.

To cope with possible nonlinearity in the data, a kernel version of the regularized MNF transformation can be obtained by reparameterization: By setting $\boldsymbol{a} \propto \boldsymbol{X}'\boldsymbol{b}$, we obtain

$$\frac{1}{NF} = \frac{\boldsymbol{b}'\boldsymbol{X}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}'\boldsymbol{b}}{\boldsymbol{b}'\left[(1-\lambda)\boldsymbol{X}\boldsymbol{X}'_N\left(\boldsymbol{X}\boldsymbol{X}'_N\right)' + \lambda\boldsymbol{X}\boldsymbol{X}'\right]\boldsymbol{b}}. \tag{3}$$

The matrices $\boldsymbol{X}\boldsymbol{X}'$ (the Gram matrix) and $\boldsymbol{X}\boldsymbol{X}'_N$ contain combinations of inner products of the rows of $\boldsymbol{X}$ only. Now, replace these inner products by inner products of nonlinear mappings of the originally measured variables into higher dimensional feature space and perform kernel substitution (the so-called kernel trick): replace the inner products by a kernel function to obtain

$$\frac{1}{NF} = \frac{\boldsymbol{b}'\boldsymbol{K}^2\boldsymbol{b}}{\boldsymbol{b}'\left[(1-\lambda)\boldsymbol{K}_N\boldsymbol{K}'_N + \lambda\boldsymbol{K}\right]\boldsymbol{b}}. \tag{4}$$

In (4), $\boldsymbol{K}_N$ is a kernelized version of the residual from a quadratic surface in a 3 by 3 window

$$\boldsymbol{K}_N = \boldsymbol{K} - \frac{-\boldsymbol{K}_1 + 2\boldsymbol{K}_2 - \boldsymbol{K}_3 + 2\boldsymbol{K}_4 + 5\boldsymbol{K}_5 + 2\boldsymbol{K}_6 - \boldsymbol{K}_7 + 2\boldsymbol{K}_8 - \boldsymbol{K}_9}{9} \tag{5}$$

where $\boldsymbol{K}$ is the usual kernel matrix between data points, and $\boldsymbol{K}_1, \ldots, \boldsymbol{K}_9$ are kernel matrices between data points in position 5 and in positions $\begin{Bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{Bmatrix}$ in the 3 by 3 neighborhood around position 5 (i.e., $\boldsymbol{K}_5$ is equal to $\boldsymbol{K}$).

We find the directions $\boldsymbol{b}$ by maximizing this Rayleigh quotient (for details, see [13], [14], [25]). kMNF variables can thus be calculated by multiplying the kernel matrix $\boldsymbol{\kappa}$ with each eigenvector from the eigenvalue problem in (4), $\boldsymbol{\kappa}$ represents a kernelization of the entire image with the training data. In the change detection procedure, each observation with the training data is kernelized respectively before the multiplication

$$\boldsymbol{kMNF}_i = \boldsymbol{\kappa}\boldsymbol{b}_i. \tag{6}$$

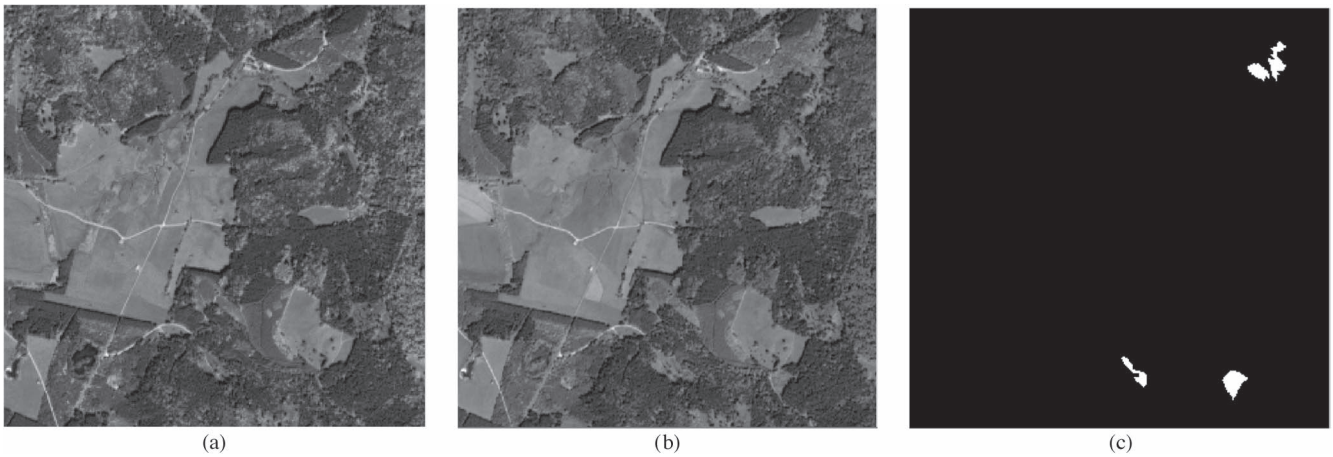

Fig. 2. Test area 1. (a) Panchromatic image of date 1. (b) Panchromatic image of date 2. (c) Digitized change reference data.

To calculate one change image based on several kMNF variates, we norm these variates to unit variance, square them, and add them up

$$\chi^2 = \sum_{i=1}^{m} \left( \frac{\boldsymbol{kMNF}_i}{\sigma_i} \right)^2. \tag{7}$$

Pixels characterized by high values are change pixels, whereas pixels with values close to zero are no-change pixels.

We use a Gaussian kernel function, and hence, we need to choose the width of the kernel $\sigma$ and the regularization parameter $\lambda$. The default choices are $\lambda = 0$ and $\sigma$ equal to the mean value of all pairwise distances between observations in the original feature space; this value is termed $\sigma_0$. In what we will call the optimized kMNF transform below, $\lambda$ and $\sigma$ are found in a simple scheme in which we use grid search to find the set of $(\sigma, \lambda)$, which maximize $1/NF$ (see also [26]). More elaborate schemes based on cross-validation can be also applied.

### C. ICDA

The idea in traditional CDA is to find projections in multi- or hypervariate feature space, which give maximal separation between groups (or classes or populations) of the data.

Here, we first use CDA with two groups based on a manually selected training area which then constitutes one of the two groups (one pixel is enough); the rest of the image is the other group. This gives rise to a potential problem: the rest of the image may contain regions that actually belong to the first group. To identify such regions and to update the training area, in a series of iterations, new training areas for the CDA are selected by automatically thresholding the canonical variate calculated in the previous iteration. Iterations stop when the canonical correlation stops improving (for details, see [21]).

## III. Experiments and Results

### A. Data Sets Description and Experiment Design

Two data sets acquired by Cartosat-1 have been selected for the experiments. The first data set was acquired over Bavaria, Germany. It exhibits an area of 2.25 km$^2$ and is shown in Fig. 2. One image is acquired in May 2008 and the other in May 2009. The region is characterized by a mixture of forest, season, and crop changes and consists of two images of size 600 $\times$ 600 pixel. In this test area, the changes of interest are related to deforestation. A change reference mask of this test site was provided by manual digitization [see Fig. 2(c)].

The second data set was acquired over Arges, Romania, near Piatra Craiului national park. Fig. 3(a) and (b) show the panchromatic image from October 2008 and November 2009, respectively. As can be seen in the panchromatic images, this region is characterized by a typical mountainous forest area, with many shadow areas caused by steep terrain. The forest changes in this test area were caused by storms; thus, within only a one-year interval, many relatively large changes can be found. In order to have a meaningful number of changes, we choose a big test site that covers an area of 20.25 km$^2$. The images are resampled to a GSD of 5 m, which leads to an image size of 900 $\times$ 900 pixel. In addition to the digitized change reference mask [Fig. 3(d)], reference vector data [Fig. 3(c)], extracted from thematic mapper data, are provided, in which only the changed regions covering an area greater than 80 000 m$^2$ are included.

As a preprocessing procedure of the experiments, for both test sites, DSMs from two dates are automatically generated with the SGM method. In addition to the 3-D coregistration of the DSMs, panchromatic images are radiometrically coregistered based on their histograms.

The proposed framework comprises two main steps: the kMNF-based change map generation and the production of the change mask based on ICDA. Therefore, two experiments have been designed in this part. In the first experiment, we generate the change maps by following the kMNF and parameter estimation approaches (hereafter referred to as kMNF-opti) described in Section II-B. The initial two-layer difference map used in this step results from the subtraction of the two panchromatic images and the two DSMs. Furthermore, the change maps resulting from image subtraction, DSM subtraction, and CVA are generated. These results are then compared with the reference change map in order to assess their accuracy.

In the first stage, to study the improvement yielded by the automatic parameter optimization method, the results are analyzed in terms of the receiver operating characteristic

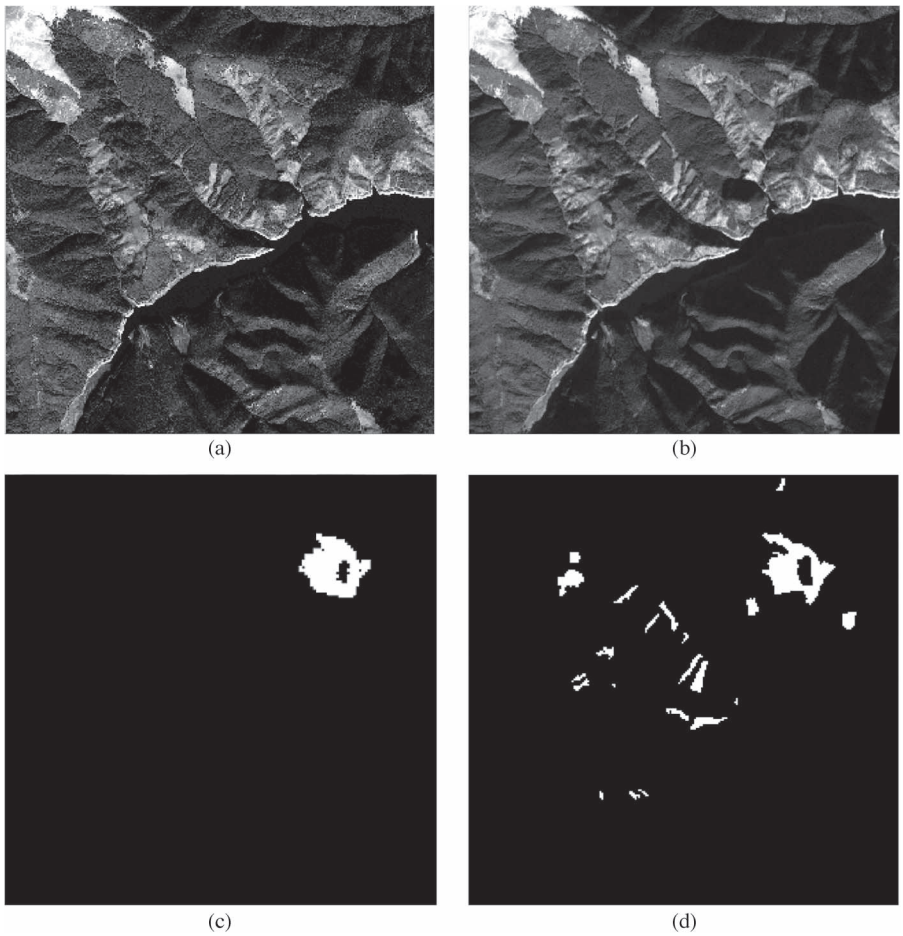


Fig. 3. Test area 2. (a) Panchromatic image of date 1. (b) Panchromatic image of date 2. (c) Reference data from vector. (d) Digitized change reference data.

(ROC) curve [27]. An ROC curve is constructed by varying the threshold from a minimum to a maximum value of the obtained change map. The change mask obtained with each threshold value can be compared with the reference data; thus, the true positive rate (Sensitivity) and false positive rate (1—Specificity) can be calculated. A larger area under the ROC curve indicates better quality of the change map.

The second experiment aims at assessing the effectiveness of the technique for labeling forest change and no-forest-change pixels. The proposed ICDA is a supervised method. As the forest-change pixels are far fewer than no-forest-change pixels, the comparison was made in terms of Kappa accuracy as the overall accuracy is not very sensitive in this case. Three other popular change labeling methods are used here for comparison. These three methods are random forests, OSVM, and k-means.

K-means [28] is an unsupervised classification method. The accuracy of k-means is closely related to the defined number of classes $k$. We set $k = 3$ in our experiments, as this value was empirically found.

OSVM: Support Vector Machine (SVM) [29], [30] is a nonlinear classification method and has been successfully used in various fields such as image classification and disease prediction. OSVM [31] tries to separate pixels belonging to one class from all other pixels. We use the OSVM implemented in LIBSVM [32], adopting a radial basis function kernel. A grid cross-validation with parameters $g$ in the range $[2^{-20}, 2^{-19.5}, 2^{-19}, \ldots, 2^{3}]$ and $v$ in the range $[0.01, 0.03, 0.05, \ldots, 1]$ are performed for each training data set. We choose one group of parameters for each size of training data.

Random forests is a robust and powerful machine learning classifier, which is capable of processing large data sets [33]. In this paper, a model with ten trees is used in the random forests classification procedure.

### *B. Change Map Generation*

The kMNF-opti illustrated in the methodology part is applied to both data sets. In our experiments, the two automatically optimized parameters for test area 1 and test area 2 are $\lambda_1 = 0$ and $\sigma_1 = 0.56\sigma_0$ and $\lambda_2 = 0$ and $\sigma_2 = 0.19\sigma_0$, respectively. The extracted change maps based on (6) are analyzed in this section. We take the first five components for the analysis in this experiment.

The change map extracted from test area 1 is shown in Fig. 4(a). It can be seen that the forest changes have been successfully highlighted in dark red. The quality improvement of the scored components from kMNF-opti can be also verified by comparing it with other change maps. The change maps used for comparison include the original kernel MNF, image subtraction, height subtraction, and CVA, which directly takes the magnitude of the height change and the image gray-value changes.

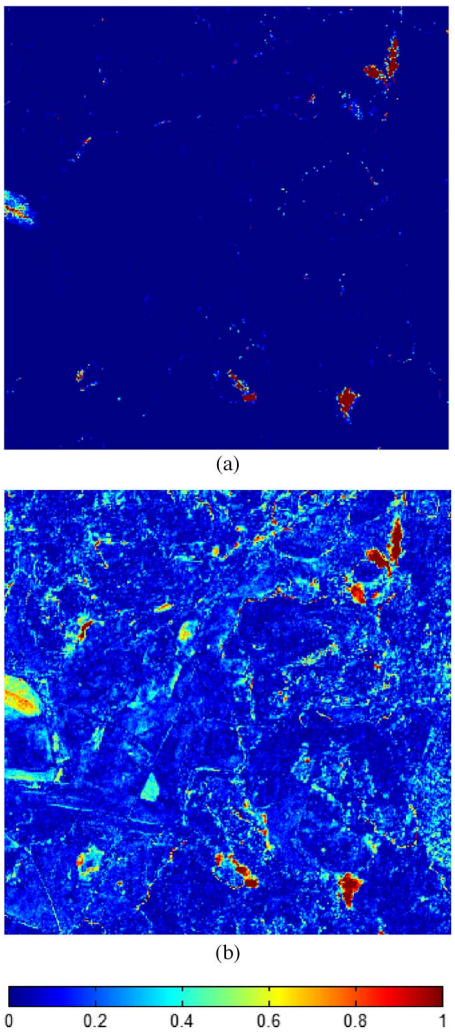


Fig. 4. Change maps of test area 1 based on (a) kMNF-opti and (b) CVA.

TABLE I
AUC COMPARISON WITH OTHER CHANGE DETECTION METHODS

| METHOD | Test area1 | Test area2 |
|---|---|---|
| ABS (IMAGE DIFFERENCE) | 0.9073 | 0.8586 |
| HEIGHT DIFFERENCE | 0.9231 | 0.9347 |
| KMNF | 0.9761 | 0.9208 |
| KMNF_OPTI | 0.9771 | 0.9251 |
| CVA | 0.9821 | 0.9441 |

kMNF-opti performs better than gray-value difference, height difference, and original kMNF in the first test area (listed in Table I). However, the CVA leads to a better ROC curve than our method. In order to prove the accuracy and stability of our method, the visual comparison of the two results is shown in Fig. 4. For a better display and comparison of these results, the 0.5% largest values are removed, and all the obtained values are linearly scaled from 0 to 1. Fig. 4 shows that the change map obtained with kMNF-opti highlights the real forest changes better than CVA.

In order to prove that the result from kMNF-opti is more capable of getting a reliable change mask, a similar method as described in [34] is adopted here. Instead of user/producer's

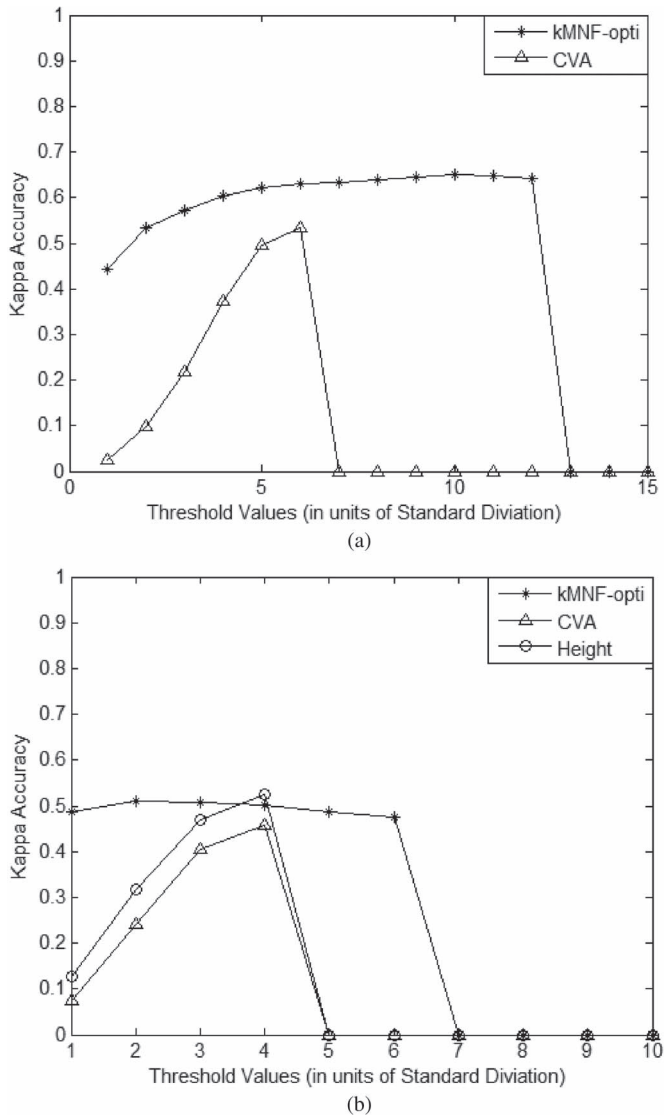


Fig. 5. Plot of the KA for (a) 15 threshold levels for test area 1 and (b) 10 threshold levels for test area 2.

accuracy, we calculate Kappa statistic (KA) in our experiment, since KA is more resistant to prevalence [35]. Threshold values are set from "1" to "15" standard deviation to produce the change mask. The obtained summary figures from these mask evaluations are plotted in Fig. 5(a). The kMNF-opti components can easily reach a KA of about 0.6, whereas the highest KA obtained from CVA is only 0.5, with higher or lower threshold values largely decreasing the change mask accuracy.

The same procedure has been applied to the second data set. According to the AUC evaluation listed in Table I, the change detection results benefit when height change information is included. The change maps obtained from kMNF-opti and CVA are shown in Fig. 6. These test images have a larger size and are located in a mountainous area; therefore, the DSM exhibits more outliers than in test area 1, and the 2% highest values are removed for better data display. From visual analyses, it is much easier to derive an appropriate threshold value for the change map in Fig. 6(a) rather than in Fig. 6(b). This is proved by the KA plot in Fig. 5(b). The height changes are also analyzed

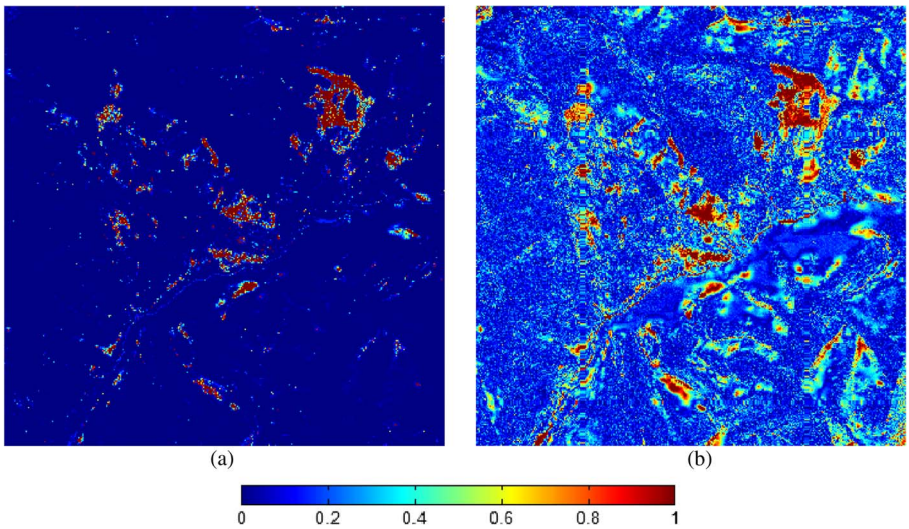


Fig. 6. Change maps of test area 2 based on (a) kMNF-opti and (b) CVA.

TABLE II
KAPPA STATISTIC COMPARISON IN TEST AREA 1

| | 1 | | 10 | | 50 | | 100 | | 200 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | mean | STD | mean | STD | mean | STD | mean | STD | mean | STD |
| ICDA | 0.4835 | 0.047 | 0.5392 | 0.029 | **0.5652** | **0.019** | 0.5578 | 0.024 | 0.5246 | 0.030 |
| K-MEANS | **0.3872** | | | | | | | | | |
| OSVM | ------ | ------ | 0.4715 | 0.038 | 0.5025 | 0.012 | 0.5024 | 0.011 | **0.5178** | **0.005** |
| RANDOM FORESTS | 0.2213 | 0.127 | **0.2355** | **0.062** | 0.1567 | 0.018 | 0.1387 | 0.025 | 0.1447 | 0.014 |
| K-MEANS* | 0.1063 | | | | | | | | | |
| OSVM* | ------ | ------ | 0.1793 | 0.165 | 0.0799 | 0.046 | **0.4236** | **0.023** | 0.4104 | 0.019 |
| RANDOM FORESTS* | **0.2130** | **0.118** | 0.2164 | 0.052 | 0.1595 | 0.037 | 0.1513 | 0.041 | 0.1432 | 0.014 |

* based on the original DSM difference and image difference. 1, 10, 50, 100, and 200 pixels are the training sizes.

here, as they exhibit a better ROC curve than kMNF-opti. The obtained KA plot is similar to the results from the first data set. In this test area, threshold values are set from "1" to "10" standard deviation, but a value of "7" is already larger than the maximum value in the change map. Although the KA value for height change is a little higher for 4 standard deviations, all the other obtained KA values are much lower than kMNF-opti. That is, it is quite difficult to set a proper threshold value for the change map from only height subtraction.

### C. Production of the Change Mask

The aim of this section is to evaluate the potential benefits of ICDA. ICDA is a supervised change labeling method. To prove the robustness of this method, results based on five groups of training sets with different sizes are analyzed. The five training sizes selected for this experiment are 1, 10, 50, 100, and 200 pixels. For each group, 14 training sets are randomly selected from the change regions. No-change training samples are additionally selected with the same procedure for those clustering methods that require this additional training, such as the random forests method.

Based on the assumption that the manually extracted change reference map represents the ground truth, KA is employed to provide evaluation results numerically. The same training data are also used for OSVM and random forests, but OSVM does not allow single-pixel values for training data. In the OSVM-based method, a cross-validation is separately processed for each training set. The average value of the calculated parameters $g$ and $v$ are used in the classification procedure. To record the result more fairly, we removed the two highest and two lowest values. The mean and standard deviation (STD) of the remaining ten values are presented in Table II and Table III. The best results with highest KA for each classifier are marked in bold.

Among these four methods, only k-means is unsupervised. ICDA and OSVM require only training samples for change regions, whereas random forests needs samples for both change and no-change areas. As listed in Tables II and III, ICDA obtains more stable and accurate results than the other two methods. The result accuracy is not strongly influenced by the size of the training data. A proper number of training samples is necessary for the OSVM classifier, as the training model cannot be built with only one sample. As can be observed, the accuracy from OSVM is improving when more training samples are used. According to our experiences, when sufficient training samples are provided, random forests can reach an accuracy value similar to ICDA. However, in this paper, random forests does not perform well due to the small size of the training sets. It has to be mentioned that the random forests classifier

TABLE III
KAPPA STATISTIC COMPARISON IN TEST AREA 2

| | 1 | | 10 | | 50 | | 100 | | 200 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | mean | STD | mean | STD | mean | STD | mean | STD | mean[%] | STD |
| ICDA | 0.4375 | 0.088 | 0.4678 | 0.006 | 0.4727 | 0.005 | 0.4811 | 0.008 | **0.4836** | **0.007** |
| K-MEANS | **0.4728** | | | | | | | | | |
| OSVM | | | 0.2254 | 0.146 | 0.4121 | 0.010 | **0.4246** | **0.005** | 0.4237 | 0.006 |
| RANDOM FORESTS | 0.2293 | 0.069 | **0.2801** | **0.065** | 0.2417 | 0.030 | 0.2425 | 0.036 | 0.2385 | 0.013 |
| K-MEANS* | **0.4677** | | | | | | | | | |
| OSVM* | ------ | ------ | 0.1390 | 0.168 | 0.2425 | 0.081 | 0.2929 | 0.012 | **0.3576** | **0.014** |
| RANDOM FORESTS* | 0.1800 | 0.155 | **0.2789** | **0.110** | 0.2225 | 0.046 | 0.2368 | 0.032 | 0.2503 | 0.018 |

* based on the original DSM difference and image difference 1, 10, 50, 100, and 200 pixels are the training sizes.

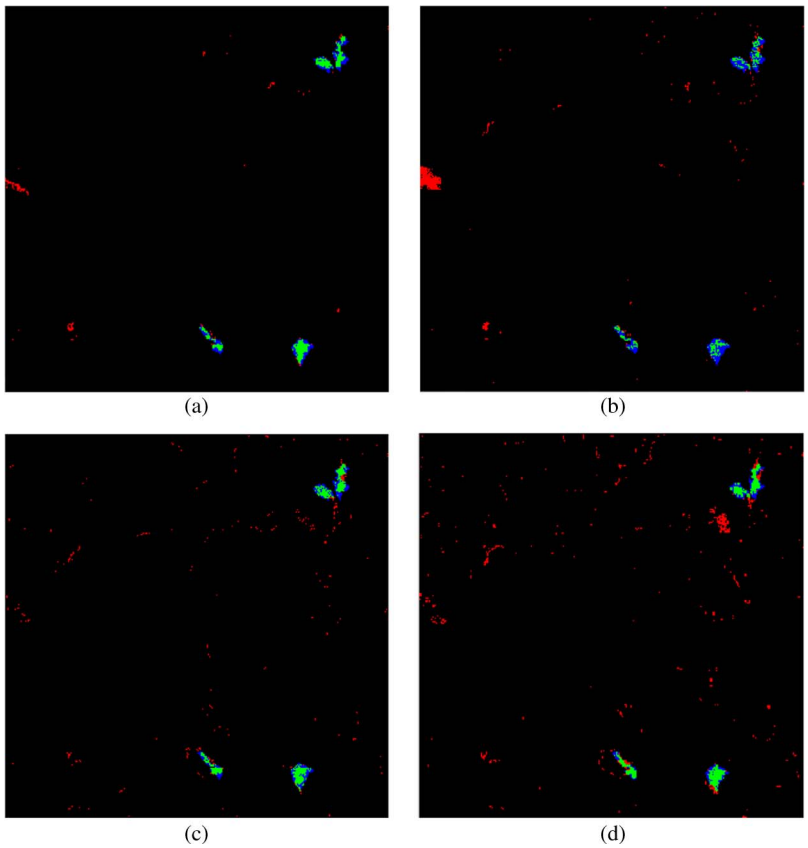


Fig. 7. Change mask from test area 1 with method (a) ICDA (KA $= 0.6216$), (b) k-means (KA $= 0.3872$), (c) OSVM (KA $= 0.5600$), and (d) random forests (KA $= 0.4955$). (Green) True detected. (Red) False alarms. (Blue) False negative.

builds decision trees randomly. Therefore, it is possible to get different classification results based on the same training data, with the numbers listed in Tables II and III being one of such test results. As our research is not focusing on random forests, no further tests are performed in this paper. It is worth noting that the results from k-means are not bad, but they require proper supervision, as it is necessary to select the $k$ parameter; furthermore, after clustering, the forest change class has to be manually selected from the available clusters.

To highlight the performance of kMNF, the accuracy values based on the raw panchromatic and DSM data are also provided in Tables II and III. These three classification methods are performed using the same training samples and approaches as transformed kMNF. As listed in Tables II and III, the results based on transformed kMNF are generally better than those yielded by the original panchromatic images and DSMs.

The obtained best change masks from each method are displayed in Figs. 7 and 8. The change mask has been overlaid to the reference change map to better illustrate the accuracy of each mask. The green color indicates the true detected pixels, the red color shows the false alarms, and the blue color gives the false negatives. Fig. 7(a) shows the change mask extracted from ICDA for the first data set. It shows more forest changes correctly extracted with less "noise". Although the other three classification methods extract most of the real forest changes, many false alarms are still present in the mask. OSVM performs

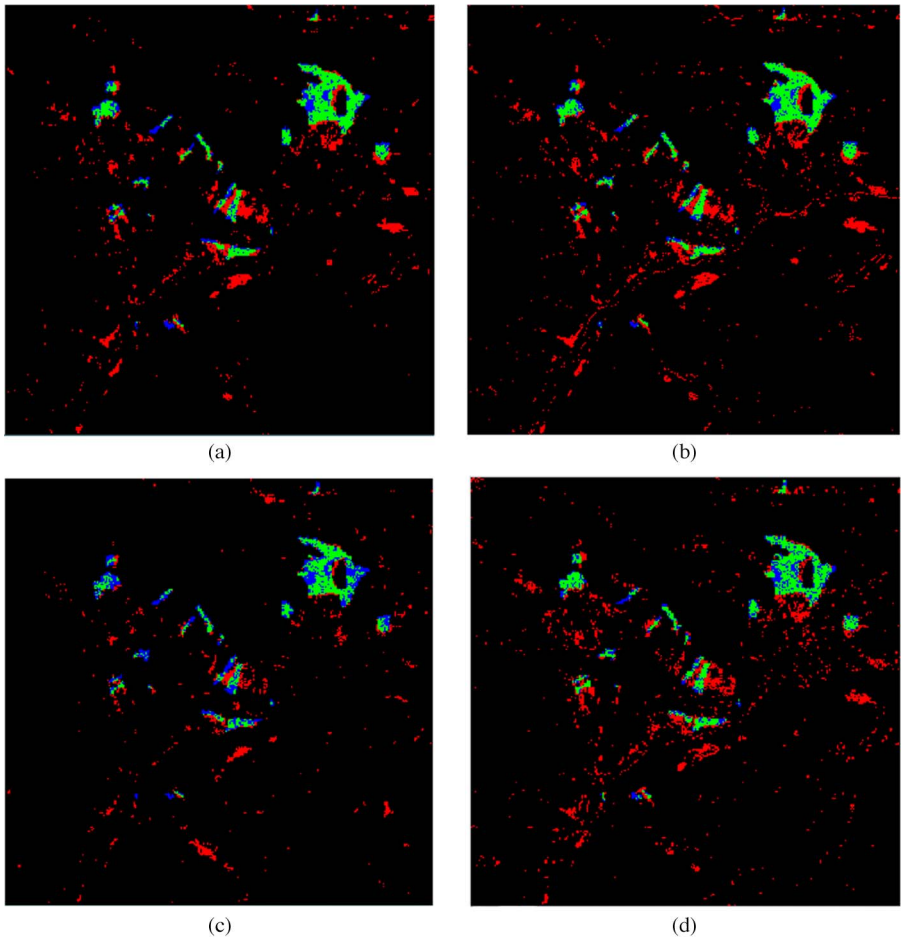


Fig. 8. Change mask from test area 2 with the methods (a) ICDA (KA $= 0.4954$), (b) k-means (KA $= 0.4728$), (c) OSVM (KA $= 0.4783$), and (d) random forests (KA $= 0.4320$). (Green) True detected. (Red) False alarms. (Blue) false negative.

only slightly worse than ICDA, but it needs the whole test region to estimate proper parameters.

The best change masks that can be extracted using these four aforementioned methods for the second data set are presented in Fig. 8. The obtained mask accuracy is not as good as in the first test area, as this second test region is more complex. In addition to the larger size and higher number of changes, this test region contains many shadow areas in the lower part, which gives problems to both automatic change detection and reference data extraction. Another reference data that could be used is the vector data provided by Humboldt University Berlin [shown in Fig. 3(c)]. OSVM can correctly label some part of this region. Random forests extracted most of this area, but exhibits large areas of false alarms, in contrast to the large number of false negative pixels from OSVM. For this area, the result from k-means is only slightly worse than ICDA.

## IV. Conclusion

In this paper, we have presented a kMNF-based forest change detection method for panchromatic stereo satellite images. First, the height map was generated based on the high-performing SGM method. Next, 2-D change features were processed onto high-dimensional feature space based on the kernel trick, and the kMNF was successfully applied to highlight the real forest changes. Finally, the ICDA was adopted to automatically extract the forest change mask.

In our methods, only two change features (height difference from DSMs and gray-value difference from panchromatic images) were used in the procedure, without any high-level change features. Furthermore, the proposed method was tested on panchromatic satellite images, which are easily available and represent relatively cheap data sources. The proposed method works well based on stereo panchromatic images and the derived DSMs. In the first step, after kMNF, changed forest areas are well highlighted in both test areas. It could be shown that it is much easier to obtain high-accuracy change masks based on the kMNF components than using other clustering methods. In the second experiment, ICDA has been compared with three other change labeling methods. Based on the same training samples, ICDA performs much better than the other methods, and the change mask accuracy is not influenced by the size of the training samples. Experimental results on the selected test data confirm that although the ICDA is a supervised change extraction method, only a small number of training pixels (even one pixel is enough) from one class (either change or no-change) are needed in order to obtain good change detection results.

The most important advantage of our method is its robustness. First, it is not restricted to specific data, although only Cartosat-1 data are tested; this method should be easily applicable to other

satellite or airborne data. This has to be proven in further investigations. If multispectral images are available, the change detection accuracy might increase even more. Second, the proposed method is not restricted to monitoring changes in forest areas. In this paper, we take forest as example, but no special texture features of forest areas are required, only DSMs and image gray-value subtraction results are needed as input. When additional training areas are available, this method can also be possibly used to highlight changes in objects such as buildings. This will be the topic of our next investigation.

## ACKNOWLEDGMENT

The authors would like to thank Euromap GmbH for providing the Cartosat-1 data for scientific purposes and Humboldt University Berlin for providing the reference vector data for test area 2. They would also like to thank Dr. D. Cerra for proofreading this paper and the anonymous reviewers and associate editors for commenting on and improving this paper.

## REFERENCES

[1] A. Singh, "Digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, Jun. 1989.
[2] J.-F. Mas, "Monitoring land-cover changes: A comparison of change detection techniques," *Int. J. Remote Sens.*, vol. 20, no. 1, pp. 139–152, Jan. 1999.
[3] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2401, Jun. 2004.
[4] S. Sader and J. Winne, "RGB-NDVI colour composites for visualizing forest change dynamics," *Int. J. Remote Sens.*, vol. 13, no. 16, pp. 3055–3067, Nov. 1992.
[5] P. R. Coppin and M. E. Bauer, "Digital change detection in forest ecosystems with remote sensing imagery," *Remote Sens. Rev.*, vol. 13, no. 3/4, pp. 207–234, Apr. 1996.
[6] R. D. Johnson and E. S. Kasischke, "Change vector analysis: A technique for the multispectral monitoring of land cover and condition," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 411–426, Jan. 1998.
[7] K. Nackaerts, K. Vaesen, B. Muys, and P. Coppin, "Comparative performance of modified change vector analysis in forest change detection," *Int. J. Remote Sens.*, vol. 26, no. 5, pp. 839–852, Mar. 2005.
[8] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.
[9] T. Celik, "Unsupervised change detection in satellite images using PCA and k-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.
[10] A. A. Nielsen, K. Conradsen, and J. J. Simpson, "Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies," *Remote Sens. Environ.*, vol. 64, no. 1, pp. 1–19, Apr. 1998.
[11] A. A. Nielsen, "The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–478, Feb. 2007.
[12] A. A. Nielsen and M. J. Canty, "Kernel principal component analysis for change detection," in *Proc. SPIE, Image Signal Process. Remote Sens. XIV*, 2008, vol. 7109, p. 10.
[13] A. A. Nielsen, "Kernel maximum autocorrelation factor and minimum noise fraction transformations," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 612–624, Mar. 2011.
[14] M. J. Canty and A. A. Nielsen, "Linear and kernel methods for multivariate change detection," *Comput. Geosci.*, vol. 38, no. 1, pp. 107–114, Jan. 2012.
[15] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised change detection with kernels," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 6, pp. 1026–1030, Nov. 2012.
[16] G. Camps-Valls, L. Gómez-Chova, J. Muñoz-Marí, J. L. Rojo-Álvarez, and M. Martínez-Ramón, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008.
[17] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[18] G. Camps-Valls and L. Bruzzone, Eds., *Kernel methods for Remote Sensing Data Analysis*. Chichester, UK: Wiley, Dec. 2009.
[19] J. Tian, P. Reinartz, P. d'Angelo, and M. Ehlers, "Region-based automatic building and forest change detection on Cartosat-1 stereo imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 79, pp. 226–239, May 2013.
[20] K. V. Kumar, T. R. Martha, and P. S. Roy, "Mapping damage in the Jammu and Kashmir caused by 8 October 2005 Mw 7.3 earthquake from the Cartosat—1 and Resourcesat—1 imagery," *Int. J. Remote Sens.*, vol. 27, no. 20, pp. 4449–4459, Oct. 2006.
[21] A. A. Nielsen and A. Müller, "Optimal class separation in hyperspectral image data: Iterated canonical discriminant analysis," in *Proc. 3rd Annu. Hyperspectral Imaging Conf.*, Rome, Italy, 2012, pp. 64–67.
[22] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 328–341, Feb. 2008.
[23] C. Straub, J. Tian, R. Seitz, and P. Reinartz, "Assessment of Cartosat-1 and WorldView-2 stereo imagery in combination with a LiDAR DTM for timber volume estimation in a highly structured forest in Germany," *Forestry*, vol. 86, no. 4, pp. 463–473, Oct. 2013.
[24] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1998.
[25] L. Gómez-Chova, A. A. Nielsen, and G. Camps-Valls, "Explicit signal to noise ratio in reproducing kernel Hilbert spaces," in *Proc. IEEE IGARSS*, Vancouver, BC, Canada, 2011, p. 3570.
[26] A. A. Nielsen and J. S. Vestergaard, "Parameter optimization in the regularized kernel minimum noise fraction transformation," in *Proc. IEEE IGARSS*, München, Germany, 2012, pp. 370–373.
[27] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine," *Clin. Chem.*, vol. 39, no. 4, pp. 561–577, Apr. 1993.
[28] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *J. R. Stat. Soc. Ser. C (Appl. Stat.)*, vol. 28, no. 1, pp. 100–108, 1979.
[29] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik, "Support vector clustering," *J. Mach. Learn. Res.*, vol. 2, pp. 125–137, Dec. 2001.
[30] B. Waske, S. Van Der Linden, J. A. Benediktsson, A. Rabe, and P. Hostert, "Sensitivity of support vector machines to random feature selection in classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2880–2889, Jul. 2010.
[31] B. Schölkopf, R. C. Williamson, A. Smola, and J. Shawe-Taylor, "Support vector method for novelty detection," in *Proc. Adv. Neural Inf. Process. Syst.*, Denver, CO, USA, 1999, vol. 12, pp. 582–588.
[32] B. Schölkopf, R. C. Williamson, A. Smola, and J. Shawe-Taylor. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/
[33] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
[34] J. T. Morisette and S. Khorram, "Accuracy assessment curves for satellite-based change detection," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 7, pp. 875–880, 2000.
[35] E. A. Freeman and G. G. Moisen, "A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa," *Ecol. Modell.*, vol. 217, no. 1/2, pp. 48–58, 2008.

**Jiaojiao Tian** received the B.S. degree in geoinformation systems from the China University of Geoscience, Beijing, China, in 2006; the M.Eng. degree in cartography and geoinformation from the Chinese Academy of Surveying and Mapping (CASM), Beijing, in 2009; and the Ph.D. degree in mathematics and computer science from Osnabrueck University, Osnabrück, Germany, in 2013.

She is currently a Research Fellow with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. Since September 2009, she has been with the "3-D group" in the Photogrammetry and Image Analysis department of IMF. In autumn 2011, she was a Guest Scientist with the Institute of Photogrammetry and Remote Sensing, ETH Zurich, Zurich, Switzerland. Her research interests include 3-D change detection for urban and forest areas, DSM quality assessment, and DSM-assisted building extraction and classification.

**Allan Aasbjerg Nielsen** received the M.Sc. degree from the Technical University of Denmark (DTU), Lyngby, Denmark, in 1978 and the Ph.D. degree from Informatics and Mathematical Modelling (IMM), DTU, in 1994.

He is currently an Associate Professor with the Department of Applied Mathematics and Computer Science, DTU. From 1977 to 1978, he was with the Danish Defense Research Establishment. From 1978 to 1985, he worked on energy conservation in housing with the Thermal Insulation Laboratory, DTU. He was with the section for image analysis from 1985 to 2001 and with the section for geoinformatics from 2001 to 2006, both at IMM. From 2007 to 2013, he was with the Danish National Space Center's section for geodesy. Since 1985, he has worked on several national and international projects on the development, implementation, and application of statistical methods, and remote sensing in mineral exploration, mapping, geology, agriculture, environmental monitoring, oceanography, geodesy, and security funded by industry, the European Union, Danida (the Danish International Development Agency), and the Danish National Research Councils.

**Peter Reinartz** (M'09) received the Dipl.-Phys. degree in theoretical physics from the University of Munich, München, Germany, in 1983 and the Dr.-Ing. degree in civil engineering from the University of Hannover, Hannover, Germany, in 1989. His dissertation is on optimization of classification methods for multispectral image data.

He is the Department Head of the department "Photogrammetry and Image Analysis" with the German Aerospace Centre (DLR), Remote Sensing Technology Institute (IMF) and holds a professorship for geoinformatics at the University of Osnabrück. He has more than 25 years of experience in image processing and remote sensing and over 200 publications in these fields. His main interests include direct georeferencing, stereo-photogrammetry and data fusion of spaceborne and airborne data, generation of digital elevation models, and interpretation of VHR data from sensors such as WorldView, GeoEye, Pleidaes a.o. He is also engaged in using remote sensing data for disaster management and using high-frequency time series of airborne image data for real-time image processing and their operational use in case of disasters as well as for traffic monitoring.