A CHANGE DETECTION METHOD FOR REMOTE SENSING IMAGE BASED ON MULTI-FEATURE DIFFERENCING KERNEL SVM

LIN Yi[a,b], LIU Bing[a,b], LV Qian-li[a,b] , PAN Chen[c], LU Yuan[a]

[a]. Research Center of Remote Sensing and Spatial Information Technology, Tongji University, Shanghai 200092, China;
[b]. Department of Surveying and Geoinformatics, Tongji University, Shanghai 200092, China;
[c]. Shanghai Municipal Institute of Surveying and Mapping, Shanghai, 200063, China

KEY WORDS: Multi-Feature Differencing Kernel; Change Detection; Support Vector Machine

ABSTRACT:

Based on the support vector machine (SVM) tools and multiple kernel method, the combinations of kernel functions were mainly discussed. The construction method of image differencing kernel with multi-feature (spectral feature and textural feature) has been developed. Through this method and weighting of the categories' samples, the improved SVM change detection model has been proposed, which could realize the direct extraction of spatial distribution information from several change classes. From the experiments we can draw the following conclusions: with the help of multiple kernel function integrating spectral features and texture information, the new change detection model can achieve higher detection accuracy than the traditional methods and is suitable for the small-sample experiment. Furthermore, it avoids the complex and uncertainty in determining change threshold required in the old detection methods.

1 Introduction

As for remote sensing image change detection, it is feasible in many application areas to construct the decision hyperplane based on single kernel function in feature space. But different feature space cannot be effectively described by single kernel function [1, 2]. In addition, it will cause a large amount of computation if classification is handled with many combined types of feature space, or if the sample feature contains heterogeneous information. Meanwhile, single-core approach is unreasonable and circumscribed in the case of uneven data distribution on high-dimensional feature space. In recent years, a new hotspot in Kernel machine learning has been developed for the above problems, i.e., kernel combination method [3]. The advantage of combining kernel function of different features to form a multi-class kernel function has higher mapping performance and more flexibility [4, 5].

With the multiple kernel method, the integration of remote sensing image multi-feature such as spectral and textural feature, an improved support vector machine (SVM) change detection model based on multi-feature differencing kernel was proposed, in which a multi-kernel function on multi-feature space was combined after applying an independent kernel function for different feature space respectively. With the new model we have obtained better results from the experiments. Fig.1 is the flow chart of Multi-Feature Differencing Kernel SVM change detection model, which will be separately discussed in detail. as follows,
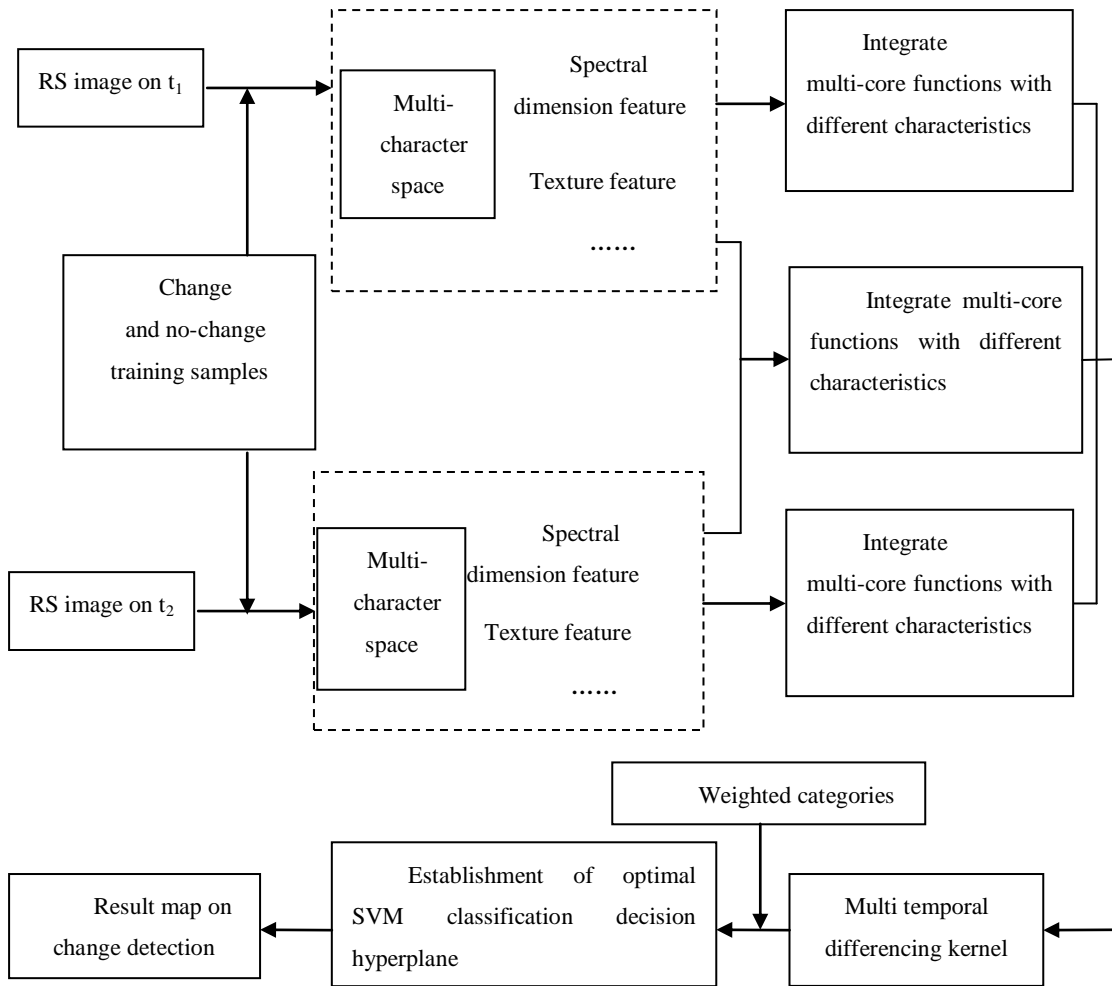
Fig.1 Flow chart of Multi-Feature Differencing Kernel SVM change detection model

2 Texture feature extraction based on gray level co-occurrence matrix (GLCM)

Gray Level Co-occurrence Matrix (GLCM) is a valid description of image texture, which is a tabulation of pixel brightness values (grey levels) occur in an image[6] . For an image I of size $M \times M$ whose gray value are i and j respectively, then the gray level co-occurrence matrix containing some certain spatial relationships is defined as the formula (1):

$$p(i,j)=\#\left\{[(r_1,c_1),(r_2,c_2)]\in M\times M\,|\,I(r_1,c_1)=i,I(r_2,c_2)=j\right\} \tag{1}$$

Where $=\#\{x\}$ denotes the number of elements in the set, (r_1, c_1) and (r_2, c_2) is a point pair on the image space, r_1,c_1,r_2,c_2 represents the pixel coordinates of the two points in the image respectively.

Let d represent the distance of the point-pair and α the direction of the two pixels, we will obtain the GLCM $p(i, j, d, \alpha)$ which varies with the distance d and the direction α. Usually, the direction α is assigned by 0 °, 45 °, 90 °and 135 °to construct 4 gray level co-occurrence matrixes while the value of pixel distance d should be decided by the nicety of the texture needed analysis.

Here in this article, five statistical quantities were chosen to realize the remote sensing image change detection [7, 8], whose formulas and physical mathematical properties are shown in the Table.1 as follows [9].

Tabel.1 Statistics of texture features

Texture Feature	Definition	Interpretation
Contrast	$$f_{con} = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1}(i-j)^2\, p(i,j)$$	Detects image edge and edge effects. For rough texture, $p(i,j)$ focuses on the main diagonal and on the contrast, it's farther away from the main diagonal.
Entropy	$$f_{ENT} = -\sum_{i=0}^{M-1}\sum_{j=0}^{M-1} p(i,j)\log p(i,j)$$	Detects complexity of the image and homogeneity of the internal space. Increases with irregularity of the texture.
Homogeneity (IDM)	$$f_{HOM} = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} p(i,j)/[1+(i-j)^2]$$	Related to contrast of the texture. Inverse difference moment, also known as the local stationary, is a measure of the rules extent of texture.
Dissimilarity	$$f_{DIS} = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1}\left|i-j\right| p(i,j)$$	$\left|i-j\right|$ represents the gray level difference of the pixel pair in scale d and the value of dissimilarity increases with the amount of pixel pairs with large gray level difference.
Correlation	$$f_{COR} = \sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\frac{i\cdot j\cdot p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y}$$	Used to determine the main direction of the texture, describe the similarity of elements in the same row or column, detects spatial variation of geometric structure of neighborhood average and variance.

Where $\mu_x = \sum_{i=0}^{M-1} i \sum_{j=0}^{M-1} p(i,j)$ $\mu_y = \sum_{j=0}^{M-1} j \sum_{i=0}^{M-1} p(i,j)$ $\sigma_x^2 = \sum_{i=0}^{M-1}(i-\mu_x)^2 \sum_{j=0}^{M-1} p(i,j)$; $\sigma_y^2 = \sum_{j=0}^{M-1}(j-\mu_y)^2 \sum_{i=0}^{M-1} p(i,j)$.

3 Generation of the spectral dimension feature set

In this paper algebraic algorithms and transformation are applied to the generation of the spectral dimension feature set for the remote sensing image change detection including the following aspects:

1) Ratio index between image bands

The band ratio can be calculated by the form [9]

$$R_{mn}(x,y) = \frac{DN_m(x,y)}{DN_n(x,y)} \qquad (2)$$

where $DN_m(x,y)$ and $DN_n(x,y)$ represent the gray level or reflection (radiation) rate in band m and n respectively in pixel (x, y) of the image.

2) Vegetation index (VI)

When we select the vegetation index to realize extraction, the vegetation types of different growth stages and varied densities should be taken into account [10].

NDVI: Normalized difference vegetation index, which is the best index factor to measure and monitor plant growth (vigor), vegetation cover and biomass production, is linear correlation to the vegetation distribution density.

DVI: Difference vegetation index

SAVI: Soil-adjusted vegetation index

GVI: Green vegetation index

GEMI: Global environment monitoring index, which varies regularly and orderly with the vegetation cycle of ups and downs in landscape environment [11].

3) Other index models

SBI: Soil brightness index

NDBI: Normalized difference building index

NWI: New water index

NDWI: Normalized difference water index

4 Construction of multi-feature differencing kernel SVM detection model

Based on the introduction and definitions about the multi-feature SVM above, with the multiple kernel method and the integration of remote sensing image multi-feature (spectral and textural feature), we could propose an improved support vector machine (SVM) change detection model, in which a multi-kernel function on multi-feature space was combined after applying an independent kernel function for different feature space respectively.

Definition 1 (Gram Matrix): Given a function $K: X \times X$, $x_1, x_2, \cdots, x_n \in X$, the element $K_{ij} = K(x_i, x_j)$ in row i, column j is the matrix named K with l rows and l columns, which is also named Gram Matrix of K and is related to $x_1, x_2, \cdots, x_n$.

Definition 2 (The Equivalent Definition of Mercer kernel):

If $K(x, x')$ is a continuously symmetric function of $X \times X$, where X represents a compact set in R^N and $K(x, x')$ is semi-positive definite about any Gram Matrix of $x_1, x_2, \cdots, x_n \in X$, $K(x, x')$ is the kernel function to meet the needs of Mercer.

Theorem: Supposed that K_1 and K_2 are kernels of $X \times X$, $X \in R^N$, and $a \geq 0$, therefore the functions as follows are both kernels.

$$K(x, x') = K_1(x, x') + K_2(x, x') \qquad (3)$$

$$K(x, x') = aK_1(x, x') \qquad (4)$$

Proof: Let K_1 and K_2 be the Gram Matrixes of K_1 and K_2 respectively relative to the finite point set $\{x_1, x_2, \cdots, x_n\}$. There exists a relational expression as the form of $\alpha^T(K_1 + K_2)\alpha = \alpha^T K_1 \alpha + \alpha^T K_2 \alpha \geq 0$ for any $\alpha \in R^N$. So $K_1 + K_2$ comes to be semi-positive definite. Hence, $K_1 + K_2$ is proved to be a kernel function through formula (3).

Meanwhile, $\alpha^T aK_1 \alpha = a\alpha^T K_1 \alpha \geq 0$, so it turns out that αK_1 is a kernel function through formula (4).

Lemma: Assumed that there are m kernels ($X \in R^N$) satisfied the Mercer conditions in $X \times X$ such as $K_1, K_2, \cdots, K_m$ and a cluster of constants labelled by β_i which is smaller than 1 but not smaller than 0 and $\Sigma_m \beta_i = 1$, we will get an equality as follows according to Theorem 1.

$$K(x, x') = \sum_{i=1}^{m} \beta_i K_i(x, x') \qquad (5)$$

The eigenvector $\mathbf{x}$ of the image were presumed to be made up of vector quantities represented by $\mathbf{x}^{(p)}$, $p=1,2,\ldots,p$, from p different information source, then the eigenvector $\mathbf{x}$ of the image can be signed by

$$\mathbf{x} = \bigcup_{p=1}^{P} \mathbf{x}^{(p)} \qquad .$$

In response to the theorem and lemma above, the compound mode of kernel function based on different information source can be divided into three categories as below [12-14]:

1) immediate accumulation kernel function

$$K(\mathbf{x}, \mathbf{x}') = \sum_{p=1}^{P} K_p(\mathbf{x}^{(p)}, \mathbf{x}'^{(p)}) \qquad (6)$$

2) weighted accumulation kernel function

$$K(\mathbf{x}, \mathbf{x}') = \sum_{p=1}^{P} \mu_p K_p(\mathbf{x}^{(p)}, \mathbf{x}'^{(p)}) \qquad (7)$$

3) intersect information kernel function

$$K(\mathbf{x}, \mathbf{x}') = \sum_{p=1}^{P} K_p(\mathbf{x}^{(p)}, \mathbf{x}'^{(p)}) + \sum_{p,p'=1}^{P} K_{p,p'}(\mathbf{x}^{(p)}, \mathbf{x}'^{(p')}) \qquad (8)$$

For multi-spectral remote sensing images, $\mathbf{x}_i$ denotes the eigenvector of each pixel in image, $\mathbf{x}_i^s$ represents the vector of texture information (spatial information) and $\mathbf{x}_i^w$ is the vector of spectral information. The specific expression of immediate accumulation kernel function and intersect information kernel function applied to the fusion of two types of feature information whose feature dimensions must be the same are given in the formula (9) and (10).

In addition, a new image differencing kernel is able to be constructed simultaneously for the purpose of change information extraction [15].

$\mathbf{x}^{(t_1)}$ and $\mathbf{x}^{(t_2)}$ are assumed respectively eigenvectors on two time phases, so for Sample i, its eigenvectors on two time phases are recorded as $\mathbf{x}_i^{(t_1)}$ and $\mathbf{x}_i^{(t_2)}$ respectively ($i=1,2,\cdots,l$, where l is the number of the sample) which are as two datasets mapped into a high-dimensional

feature space in a nonlinear mode, where linear learning is realized .The difference of the samples on two time

(11). Then Sample i turns to be $\phi_i(\mathbf{x}) = \phi_i(\mathbf{x}^{(t1)}) - \phi_i(\mathbf{x}^{(t2)})$.

At the same time, the corresponding dot product of the kernel function can be expressed as the formula (12)[16].In that case, the difference kernel function is in the form of the formula (13).

On the understanding that the sample feature at different phases contains spectrum information as well as texture information [17- 20], it is necessary to apply intersect

phases as a new input sample is calculated as the formula

information kernel function into the multi-feature information in the first place and then a difference kernel at varied time phases can be constructed on this basis[21]. As a result the difference kernel function of different time phases based on multi-feature space comes out as follows the formula (14):

$$K(\mathbf{x}_i, \mathbf{x}_j) = K_s(\mathbf{x}_i^s, \mathbf{x}_j^s) + K_w(\mathbf{x}_i^w, \mathbf{x}_j^w) \tag{9}$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = K_s(\mathbf{x}_i^s, \mathbf{x}_j^s) + K_w(\mathbf{x}_i^w, \mathbf{x}_j^w) + K_{sw}(\mathbf{x}_i^s, \mathbf{x}_j^w) + K_{ws}(\mathbf{x}_i^w, \mathbf{x}_j^s) \tag{10}$$

$$\begin{aligned} \varphi(\mathbf{x}) &= \{\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \cdots\} \\ &= \varphi(\mathbf{x}^{(t1)}) - \varphi(\mathbf{x}^{(t2)}) \\ &= \{\phi_1(\mathbf{x}^{(t1)}) - \phi_1(\mathbf{x}^{(t2)}), \phi_2(\mathbf{x}^{(t1)}) - \phi_2(\mathbf{x}^{(t2)}), \cdots\} \end{aligned} \tag{11}$$

$$\begin{aligned} K(\mathbf{x}, \mathbf{x}') &= \varphi(\mathbf{x}) \cdot \varphi(\mathbf{x}') = (\varphi(\mathbf{x}^{(t1)}) - \varphi(\mathbf{x}^{(t2)})) \cdot (\varphi(\mathbf{x}'^{(t1)}) - \varphi(\mathbf{x}'^{(t2)})) \\ &= \varphi(\mathbf{x}^{(t1)}) \cdot \varphi(\mathbf{x}'^{(t1)}) + \varphi(\mathbf{x}^{(t2)}) \cdot \varphi(\mathbf{x}'^{(t2)}) - \varphi(\mathbf{x}^{(t1)}) \cdot \varphi(\mathbf{x}'^{(t2)}) - \varphi(\mathbf{x}^{(t2)}) \cdot \varphi(\mathbf{x}'^{(t1)}) \end{aligned} \tag{12}$$

$$K(\mathbf{x}, \mathbf{x}') = K(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t1)}) + K(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t2)}) - K(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t2)}) - K(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t1)}) \tag{13}$$

$$\begin{cases} K(\mathbf{x}, \mathbf{x}') = K(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t1)}) + K(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t2)}) - K(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t2)}) - K(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t1)}) \\ K(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t1)}) = K_s(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t1)}) + K_w(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t1)}) + K_{sw}(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t1)}) + K_{ws}(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t1)}) \\ K(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t2)}) = K_s(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t2)}) + K_w(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t2)}) + K_{sw}(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t2)}) + K_{ws}(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t2)}) \\ K(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t2)}) = K_s(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t2)}) + K_w(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t2)}) + K_{sw}(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t2)}) + K_{ws}(\mathbf{x}^{(t1)}, \mathbf{x}'^{(t2)}) \\ K(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t1)}) = K_s(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t1)}) + K_w(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t1)}) + K_{sw}(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t1)}) + K_{ws}(\mathbf{x}^{(t2)}, \mathbf{x}'^{(t1)}) \end{cases} \tag{14}$$

Confessedly, the general expression for the solution to SVM optimization in dual form can be described as:

$$\begin{cases} \max: L(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\ S.T. \qquad C \geq \alpha_i \geq 0, i = 1,2,\cdots,l \\ \qquad \sum_{i=1}^{l} \alpha_i y_i = 0 \end{cases}$$

$$\begin{cases} \max: L(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j) = \sum_{i=1}^{l} \alpha_i - \\ \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j [K(x_i^{(t1)}, x_j^{(t1)}) + K(x_i^{(t2)}, x_j^{(t2)}) - K(x_i^{(t1)}, x_j^{(t2)}) - K(x_i^{(t2)}, x_j^{(t1)})] \\ S.T. \qquad C \geq \alpha_i \geq 0, i = 1,2,\cdots,l . \\ \qquad \sum_{i=1}^{l} \alpha_i y_i = 0 \end{cases} \tag{15}$$

Thereupon, put the formula (14) into the one above and then we come at the multi-feature difference kernel SVM change detection model based on multi-feature

difference kernel function in the form of the formula (15).

The combination of the multi-feature multi-kernel function and multi-phase difference kernel puts the immediate change information extraction into practice, bring the integration of a wide variety of information sources into effect and avoid the complexity and contingency of selecting threshold as well [23, 24].

5 Experimental results and analysis

We choose TM multispectral images on band 1~5 and band 6 of some certain islands in Yangtze River Delta Area taken on April, 25 of 2008 and July, 17 of 2009, in which the square measure of the experimental area is 5625 hectares. Compared with the land-use map, the main surface features in this experimental area are divided into five categories, signed water as C1, structures C2, buildings C3, vegetation C4 and bare ground C5[25]. The original remote seining images at the band543-RGB from 2008 and 2009 are presented in Fig.2. These experiments will be analyzed in detail as follows.
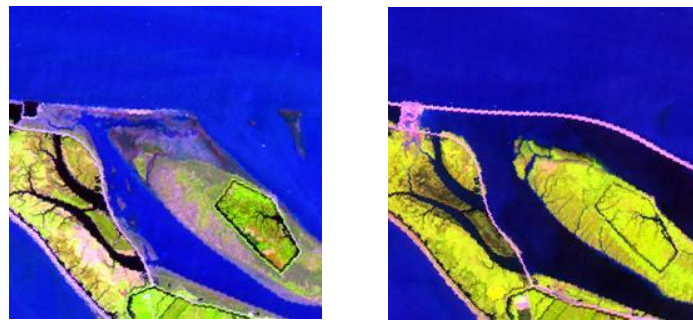

(a) original image of 2008 (band 543-RGB)   (b) original image of 2009 (band543-RGB)

Fig. 2 Remote sensing images of the experimental area

1) Through the choice of each sample extracted from change and non-change areas with semi-automatic human-computer interaction method, the dimension of multi-feature on two time phases is 24 in total. We select 1472 samples totally including 532 samples from change areas and 940 samples from non-change areas, of which 60% are selected randomly as training samples and the other as test samples. Accordingly, the weight of changed category and non-changed one are calculated as 0.64 and 0.36, respectively.

2) Multi-feature information of images on every time phase including 6 dimension spectral feature and 6 dimension texture feature is extracted in method discussed in section 1 and 2. Fig.3 shows the texture images on band 3, 4 and 5, in which the image from 2009 is taken as example.
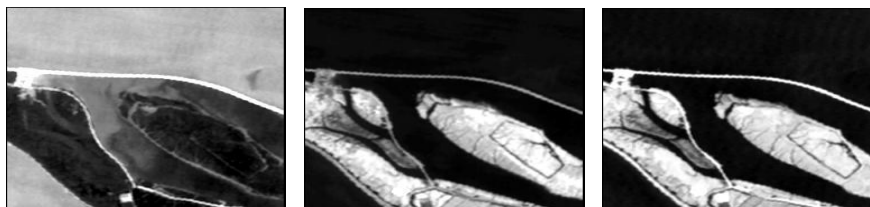
(a) band 3   (b) band 4   (c) band 5

Fig. 3 Textural extraction results of 2009 remote sensing image (band 3, 4 and 5)

3) The intersect information kernel on and between different time phases can be acquired according to the combination model of the formula (10).

4) According to formula (13), we begin with the construction of multi-feature difference kernel, where the kernel functions of each part were obtained through the intersect information kernel achieved from the processes above and then we set up the optimal separating hyperplane on the basis of the algorithm of multi-feature difference kernel based SVM. After the treatment above,

we gained the change detection results of the experimental areas from 2008 to 2009 as shown in Fig.4 (d) .

In order to certificate the change detection model of multi-feature difference kernel based on SVM, three representative kinds of change detection methods, the change detection method in difference image on a single band with the highest precision (band 2), the first principal components change detection method and the method of simultaneous classification on multi-phases, were applied into the same test samples. Fig.4 (a), (b), (c) show the change detection results of the experimental areas as follows.

In Fig.4, (a) shows the result of the change detection method in difference image on a single band with the highest precision (band 2); (b) shows the result of the first principal components change detection method; (c) shows the result of the method of simultaneous classification on multi-phases; (d) shows the result of the multi-feature difference kernel change detection based on SVM.

It can be seen from Fig.4 that in image (a) and (b), there exist dramatic noise and a loss of edge information, which is obviously shown in the change area of ditch on upper side of the image. From image (c), we can find the method applied here better avoided the effect of noise, but mistook some vegetation area influenced by the radiation difference for change area because of the radiation difference of varied phase images.

The comparison result of quantitative analysis was listed in the Table.2 below.
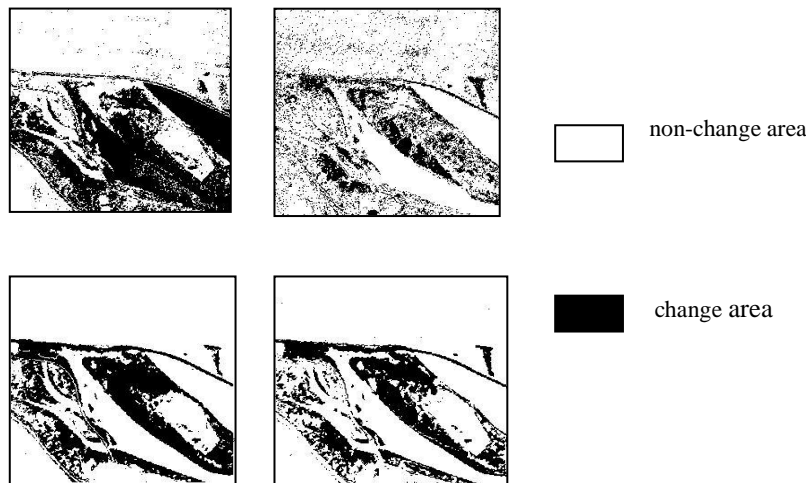

Fig. 4 The change detection results of the experimental area.

The characteristics of the change detection model of multi-feature differencing kernel based on weighted SVM can be drawn from these experiments as follows:

1) The indexes of overall detection accuracy and Kappa coefficient calculated as 91.0% and 0.8045 respectively in the method of multi-feature differencing kernel based on weighted SVM are larger than that in the other three methods. This method is able to realize the combination of the change information in varied images from all bands and a variety of features, which overcomes the limitations of a single band and the lack of information and can achieve more complete and accurate results.

Tabel.2 Accuracy of four change detection methods

Method	Probability of detection (%)	Probability of miss (%)	Probability of false positive (%)	Overall detection accuracy (%)	Kappa coefficient
Single-band difference method	44.3	31.2	24.5	73.3	0.4113
Difference between the principal component analysis	36.0	35.0	29.0	67.4	0.2808
Multi-temporal simultaneously classification method	69.9	7.2	22.8	85.9	0.7076
Multi-feature differencing kernel SVM detection method	77.7	11.8	10.5	91.0	0.8045

2) Although Multi-temporal simultaneous classification method can also be seen as a combination of multi-kernel functions, it is just a simple summation of the different features and different phase. While the method of Multi-feature differencing kernel based on SVM a) applies the combined kernel functions, in which the data from different feature space is mapped through inputting into the corresponding kernel function, and then b) construct the multi-feature differencing kernel. These made the data in a new feature space be better expressed with effective fusion of information on different time phase, which are supported by higher level of the four accuracy indexes (i.e. overall detection accuracy, probability of detection, probability of miss and probability of false positive).

3) The change results can be direct detected by multi-feature differencing kernel based on SVM, without the procedure for setting a threshold accompanied with redundant artificial interference. This leaves out the complexity and uncertainty in searching optimal threshold.

Reference

[1] PAN Chen, LIN Yi and CHEN Ying-ying. Decision tree classification of remote sensing images based on multi-feature [J]. Journal of Optoelectronics Laser, May, 2010, 21(5):731-736. (In Chinese)

[2] WANG Hong-qiao, SUN Fu-chun, CAI Yan-ning, CHEN Ning and DING Lin-ge. On multiple kernels learning methods [J], ACTA AUTOMATICA SINICA, August, 2010, 36(8):1037-1048. (In Chinese)

[3] Bach F R. Consistency of the group Lasso and multiple kernels learning [J]. The Journal of Machine Learning Research, 2008, 9(6):1179-1225.

[4] Ozogur-Akyuz S, Weber G W. Learning with infinitely many kernels via semi-infinite programming. In: Proceedings of the EURO Mini Conference on Continous Optimization and Knowledge Based Technologies. Neringa, Lithunia: Vilnius Gediminas Technical University Publishing House, 2008, 342-348.

[5] Damoulas T and Girolami M A. Pattern recognition with a Bayesian kernel combination machine. Pattern Recognition Letters, 2009, 30(1):46-54.

[6] Vedaldi A, Gulshan V, Varma M, Zisserman A. Multiple kernels for object detection. In

Proceedings of the International Conference on Computer Vision. Tokyo, Japan: IEEE, 2009, 1-8.

[7]   Wang Z, Chen S, SUN T. MultiK-MHKS: a novel multiple kernel learning algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2)348-353.

[8]   Manik V, Rakesh B. More generality in efficient multiple kernel learning. In: Proceedings of the 26th International conference on Machine Learning. Monthreal, Canada: ACM, 2009, 1065-1072.

[9]   Tang L, Chen J, Ye J. On multiple kernel learning with multiple labels. In Proceedings of the 21st   International Joint Conference on Artificial Intelligence. Pasadena, USA: Morgan Kaufmann Publishers, 2009, 1255-1260.

[10] Kembhavi A, Siddiquie B, Miezianko R, McCloskey S, Davis L S. Incremental multiple kernel learning for object recognition. In: Proceedings of the International Conference on Computer Vision. Washington D.C., USA: IEEE, 2009, 1-8.

[11] Zhao B, Kwok J T, Zhang C S. Multiple kernel clustering. In: Proceedings of the 9th SIAM International Conference on Data Mining. Philadelphia, USA: SIAM, 2009, 638-649.

[12] Mu Shao-min, Tian Sheng-feng, Yin Chuan-huan. Multiple kernel learning based on cooperative clustering. Journal of Beijing Jiaotong University, 2008, 32(2):10-13. (In Chinese)

[13] Chi-Yuan Yeh, Wen-Pin Su, Shie-Jue Lee. Improving efficiency of multi-kernel learning for support vector machines. In: Machine Learning and Cybernetics, 2008 International Conference on Data Mining, China:Kunming, 2008,12(15), 3985-3990.

[14] Subrahmanya N, Shin Y C. Sparse multiple kernel learning for signal processing applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(5):788-798.

[15] Gustavo Camps-Valls, Luis Gómez-Chova, Jordi Muñoz-Marí etal . Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection．IEEE Transactions on Geoscience and Remote Sensing, 2008, 46(6).

[16] LIU Li, KUANG Gang-yao. Overview of Image Texture Feature Extraction Methods[J]. Journal of Image and Graphics, April, 2009,14(4):622-635. (In Chinese)

[17] HE Wei-zhong, ZHOU Wei-na, SU Ping. Research on classification and target recognition of remote sensing image based on improved support vector machine [J]. Remote Sensing Information, 2010(6):6-8.(In Chinese)

[18] Guo Lei, Xiao Huai-tie and Fu Qiang. The study on the separability of SVM used in target recognition [J]. Journal of Electronics and Information Technology, March, 2009, 31(3):570-573.(In Chinese)

[19] Allan Aasbjerg Nielsen, Morton John Canty. Kernel principal component analysis for change detection.In: SPIE Europe Remote Sensing Conference on Geoinformatics,Sep.2008,vol.7109.

[20] Tuia, D., Camps-Valls, G., Matasci, G., Kanevski, M.. Learning relevant image features with multiple-kernel classification. IEEE Transactions on Geoscience and Remote Sensing, Oct., 2010, 48(10): 3780-3791. [21] M. Fauvel, J. A. Benediktsson, J. Chanussot and J. R. Sveinsson. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles, IEEE Trans. Geosci. Remote Sens., 2008(46): 3804.

[22] G. Camps-Valls, L. Gmez-Chova, J. Muoz-Mar, J. L. Rojo-lvarez and M. Martnez-Ramn   Kernel-based framework for multi-temporal and multi-source remote sensing data classification and change detection. IEEE Trans. Geosci. Remote Sens., 2008(46):1822.

[23] B. Demir and S. Ertrk. Phase correlation based redundancy removal in feature weighting band selection for hyperspectral images.   Int. J. Remote Sens., 2008(29):1801.

[24] B.-C. Kuo, C.-H. Li and J.-M. Yang. Kernel nonparametric weighted feature extraction for hyperspectral image classification.   IEEE Trans. Geosci. Remote Sens.,   2009(47):1139. [25] D. Tuia, F. Ratle, A. Pozdnoukhov and G. Camps-Valls . Multi-source composite kernels for urban image classification. IEEE Geosci. Remote Sens. Lett., 2010(7):88.