Change Field: A New Change Measure for VHR Images

Leigang Huo, Xiangchu Feng, Chunlei Huo, *Member, IEEE*, Zhixin Zhou, and Chunhong Pan

***Abstract*—Due to the complexity of very high resolution (VHR) images and the inaccurate correspondence, change feature extraction is the key difficulty of VHR image change detection. In this letter, change field is proposed to represent the complex changes between VHR images. Change field measures the complex changes based on the displacements and the compensated distance. Based on change field, a novel change detection approach is proposed, where the inter-class variability is improved and the changed class and the unchanged class can be separated effectively. Experiments demonstrate the effectiveness of the proposed approach.**

***Index Terms*—Change detection, change field, progressive transductive classification, very high resolution (VHR) images.**

I. Introduction

COMPARED to the traditional low-to-moderate resolution images, VHR image change detection is more attractive. Besides the location where changes occur, the types and the details of the changes can also be recognized by taking advantages of the improved spatial resolution. However, new challenges are raised simultaneously, among which change feature representation (i.e., how to represent the complex changes between VHR images. This topic will be illustrated in depth in Section II) is the key factor that suspends the practical application. In detail, the main challenges lie in the following two aspects.

First, the discriminability between different land-cover classes is determined simultaneously by the spatial resolution and the spectral resolution. Although the spatial resolution of VHR sensors increases significantly, the spectral resolution and the spectral separability are relatively poor [1]. As a result, the statistical separability of the different land-cover classes is reduced in the spectral domain. For the similar reason, the overall separability of the changed class and the unchanged class presented in VHR images is not improved with the increased spatial resolution.

Manuscript received July 15, 2013; revised November 26, 2013 and January 29, 2014; accepted February 26, 2014. Date of publication March 19, 2014; date of current version May 12, 2014. This work was supported by the Natural Science Foundation of China under Grants 61271294, 61375024, 61005013, 91338202, 61105011 and 863 project 2013AA7013031.

L. Huo is with the School of Mathematics and Statistics, Xidian University, Xi'an 710071, China and also with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: lghuo2@126.com).

X. Feng is with the School of Mathematics and Statistics, Xidian University, Xi'an 710071, China (e-mail: xcfeng@mail.xidian.edu.cn).

C. Huo and C. Pan are with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: clhuo@nlpr.ia.ac.cn; chpan@nlpr.ia.ac.cn).

Z. Zhou is with the Beijing Institute of Remote Sensing, Beijing 100191, China (e-mail: zhixin.zhou@mail.ia.ac.cn).

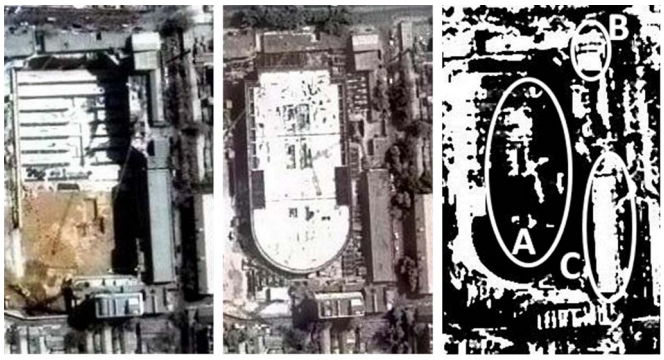
Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2014.2310202

Second, the overall separability is further deteriorated by the inaccurate correspondence. Generally, the inaccurate correspondence is mainly caused by the image registration error, view angle variation, etc. For instance, many uncertainties (such as feature matching, outlier removal, transformation model selection and estimation, interpolation artifact, pan-sharpening error) are involved in image registration procedure, and the registration error is accumulated by the above impacts. Accurate VHR image registration is still an open problem, and it hinders the reliable change detection. For another example, the appearances and the positions of the same object(e.g., the buildings) varies greatly with the view angle difference, and it is difficult for the traditional global-transformation-based image registration approaches to compensate the nonlinear deformation. For low-to-moderate resolution images, the above impacts can be neglected, and the satisfactory performance can be achieved by the spectral differences. However, for VHR images, many false changes are produced and difficult to remove due to the inaccurate correspondence.

In short, the complexities of the changes lie in the fact that more false negatives and false positives are generated by the improved spatial resolution and difficult to be removed. Many change features proposed in the existing literature aimed at addressing the above difficulties by: 1) improving the discriminativeness of the features to effectively represent the complex objects in an image or 2) improving the robustness of the change features to reduce the undesired impacts. For instance, Huo [2] proposed to extract object-specific change feature which combines the relative change feature and the original multi-temporal signatures. Falco [3] proposed to use morphological attribute profiles to represent multiscale geometrical features, which is implemented by multiresolution contextual transformation. To remove the false changes caused by the registration noise, Bovolo [4] proposed to represent change features by spectral change vectors in the polar domain. To reduce the parallax effects caused by the camera motion, Bourdis [5] proposed to measure the change by the residual error, which is computed based on the constrained optical flow matching. Similarly, Klaric [6] presented to calculate the "corrected" feature difference based on the corrected pixel correspondence and the window-based bidirectional difference algorithm.

Despite of the novelties of the above change metrics, they are lack of meaningful interpretation and insufficient to measure the complex changes. In this context, it is necessary to develop new change metrics for VHR images. To address the above difficulties, a novel change measure called change field is proposed for VHR images. Compared to the traditional metrics, change field is promising in utilizing high-dimension local



(a) (b) (c)

Fig. 1. Illustration of shortcomings of traditional change features based on pixel-wise difference. (a) and (b): multi-temporal images, (c): change map based on pixel-wise spectral difference.

features to represent the complex changes and compensating the inaccurate correspondence.

II. Change Field

Considering the shortcomings of the traditional pixel-wise difference approaches, change field aims at capturing the complex changes between VHR images via simulating the visual change detection of human beings. For this reason, we investigate the behaviors of the traditional approaches on VHR images and the complex changes caused by the improved spatial resolution. As illustrated in Fig. 1, for a co-registered VHR image pair, change detection errors caused by the pixel-wise differencing approach can be generally classified into the following two types:

1) False negative. As shown in region A in Fig. 1, false negatives are mainly due to the low interclass variability(different objects are of similar spectral features, for instance, the road and the building roof) and the limitation of the pure usage of the spectral features. In fact, it is difficult to represent such complex changes even with the help of multiband spectral features if only the spectral feature is used. In other words, higher level features are important to reduce missed alarms.

2) False positive. As illustrated in region B and C in Fig. 1, most false positives are caused by the registration error and view angle variation. As reported in [7], sub-pixel image registration accuracy is needed for change detection. However, this accuracy cannot be achieved at the building areas, the underlying reason lies in the facts that the traditional image registration approaches utilize the global transformation model and no additional compensations are considered for the nonlinear deformation. In consequence, for the building regions under different view angles, it is impossible to avoid high false positives even if discriminative features are used.

Surprisingly, human beings can detect the changes robustly to the above impacts. Human eyes will subconsciously encode the complex objects from the spectral space into the feature space and search the similar objects in the other image within certain ranges before comparison and decision. In other words, human beings decide whether changes occur between objects based on the differences after compensation [8].

Motivated by the human visual change detection mechanism, we define **change field** as follows:

For the co-registered multi-temporal images $I1$ and $I2$, **change field** is a 3-D vector (x, y, z) defined over each pixel (i, j) that follows the two conditions

$$(x,y)=arg\min_{x_i,y_j}\sum_{\substack{i,x_i\in N(i),\\ j,y_j\in N(j)}} D(F(I_1(i,j)),F(I_2(i+x_i,j+y_j))) \tag{1}$$

$$z=D\left(F\left(I_1(i,j)\right),F\left(I_2(i+x,j+y)\right)\right) \tag{2}$$

where F and D denote the discriminative feature space (e.g., SIFT [9], DAISY [10]) and the robust distance function (e.g., Mahalanobis distance and Earth Mover's distance) respectively, $N(i)$ and $N(j)$ are the neighborhoods for the pixel indices i and j. As can be seen from the above definition, (x, y) is the displacement vector and aims at simulating human eyes to find the local search range that makes the similarity between the patch pairs in I_1 and I_2 maximized. z is the distance after compensating the necessary displacements. For a pixel or object p in I_1 and its counterpart q in I_2, the direct interpretation of change field is that p reaches q after the displacement vector (x, y) with the cost z.

Given the change field concept, the collection of (x, y, z) over all pixels lies within the following semi-ellipsoid [Fig. 2(a)]

$$\frac{(x-C_x)^2}{R_x^2}+\frac{(y-C_y)^2}{R_y^2}+\frac{(z-C_z)^2}{R_z^2}\le 1, z\ge 0 \tag{3}$$

where R_x, R_y, R_z denote the maximum of x, y, z along the directions of X, Y and Z, respectively. (C_x, C_y, C_z) represents the clustering center of the unchanged class. In the ideal case, the center of (C_x, C_y, C_z) should be the coordinate origin, i.e., $(C_x, C_y, C_z) = (0, 0, 0)$. In the practical situations, due to the impacts caused by the mis-registration or view-angle variation, (C_x, C_y, C_z) will violate the coordinate origin. As illustrated by Fig. 2(b), the semi-ellipsoid is divided into three layers: the unchanged class (green), the false changed class (blue), the changed class (red). By change field, change features can be distinguished effectively: the unchanged class is characteristic of small displacement and cost, the changed class is characteristic of huge displacement or cost, while the false changed class corresponds to small cost and reasonable displacement. Considering the similarities of the unchanged class and the false changed class in the spatial distribution and the semantic interpretation, we combine the false changed class and the unchanged class.

To clarify the novelty of the change field concept, we analyze the similarities and the differences between some related concepts.

1) Change field is similar to change vector [4] in projecting high-dimension feature space into low-dimension change feature space. However, change vector [4] is 2-D, and change features are obtained by the pixel-wise algebra operation. In contrast, change field is 3-D, and it compensates the local non-rigid displacements adaptively.

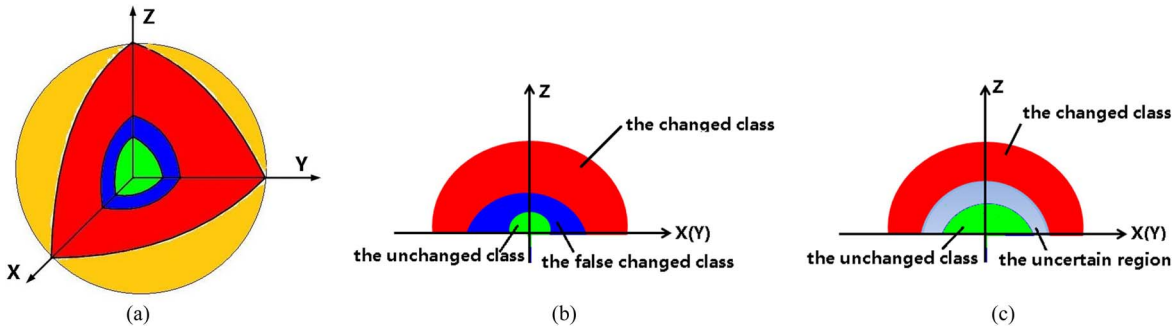


Fig. 2. Illustration of change field. (a): 3-D change field, (b) 2-D cross section of 3-D change field in ideal case. (c) Change field distribution in real case.

2) Change field is similar to optical flow and SIFT flow in containing the displacement information. However, the differences between them lie in many aspects. First, optical flow or SIFT flow focuses mainly on the 2-D motion information, while change field aims at capturing the change signature. Of course, changes are closely related to the motion [8], but the motion is only one of the signatures for changes. For the co-registered multi-temporal images, the item (x, y) in change field is not for the motion information but to compensate the registration error or view angle variation. Second, change field is a 3-D vector. It contains not only the displacement vector, but the change probability after compensation, i.e., the cost z combined with (x, y) used to separating the changed class from the unchanged class. Huo [11] computes the displacements based on SIFT flow and uses only the displacements as the change features. However, as illustrated in the experiment section, the cost z is helpful in separating the false changes from the real changes.

III. Change Field Analysis

Change field is not a specific change detection approach but a general change metric for VHR images. To illustrate the effectiveness of change field, a novel change detection approach called change field analysis is proposed in this letter, which consists of the following two steps: change field computation and change field classification.

A. Change Field Computation

As stated before, discriminative features are required to improve the interclass variability. Compared to the spectral features, local feature descriptors (e.g., HOG or SIFT) are promising in capturing the complex structure of the objects, and they are used widely in image matching, image classification and object recognition [12]. Winder [13] has recently shown the superiority of the DAISY descriptor [10] in comparison to other widely used descriptors such as SIFT. For this reason, in this letter, DAISY is used for dense descriptor, i.e., at each pixel, we use DAISY to encode the local neighborhood.

After DAISY features are extracted at each pixel from the images individually, change field can be computed based on feature matching. Due to the repetitive structures contained in VHR images, the traditional global feature matching such as KNN is not suitable for our purpose. Instead, local nearest neighbor search is used for change field computation. For a pixel $p_1 = (i, j)$ from the image I_1, local nearest neighbor is the pixel $q_1 = (k, l)$ in the image I_2 that makes the distances of DAISY features at p_1 from I_1 and the neighbors around p_1 in I_2 minimized. Change field at p_1 is $(x, y, z) = (|i - k|, |j - l|, dis(p_1, q_1))$, where $dis(p_1, q_1)$ is the Euclidean distance between DAISY features extracted at p_1 and q_1. To constrain the spatial smoothness and speed up the computation, a fast deformable spatial pyramid matching algorithm [14] is used to find the nearest neighbor for each pixel and compute the change field. The size of the neighborhood is determined by the satellite acquisition geometry, the spatial resolution, etc. For QuickBird images used in the experiments, we found that the satisfactory performances can be achieved when the size of the neighborhood is in the range [30,60].

B. Change Field Classification

The distribution of change field is illustrated in Fig. 2(b). The interclass between the changed class and the unchanged class (including the false changed class) is improved by change field, and change map can be achieved by the traditional classification approaches in a supervised or unsupervised fashion. To balance the laborious procedure in labeling the training samples and the high classification accuracy, we utilize a progressive semi-supervised classifier [15].

For simplicity, we use $l_i = -1$, 1 to denote the label of the unchanged class and the changed class respectively. $l_i = 0$ denotes the label to be determined for the test samples. Based on the training samples $tr_i = (f_i, l_i)$ and the trained classifier $F(\cdot)$, the labels of the test samples, te_i can be determined, i.e., $l_i = F(f_i)$, where f_i is the change field at the i-th pixel. It is difficult and impractical to select the representative training samples automatically. As illustrated by Fig. 2(c), based on the spatial distribution, change field can be divided into three levels: the changed class, the unchanged class(including the false changed class) and the uncertain region. The uncertain region means the region where the changed class is mixed up with the unchanged class. The principle of progressive semi-supervised classifier is to classify the change features in the uncertain region progressively driven by the training samples of high confidence. The confidence is measured by the distance of change field to the coordinate origin. For the unchanged class, the smaller the distance is, the higher the confidence will be. For the changed class, the bigger the distance is, the higher the confidence will be. By this way, the classifier is initially determined by the above samples, and it is refined by

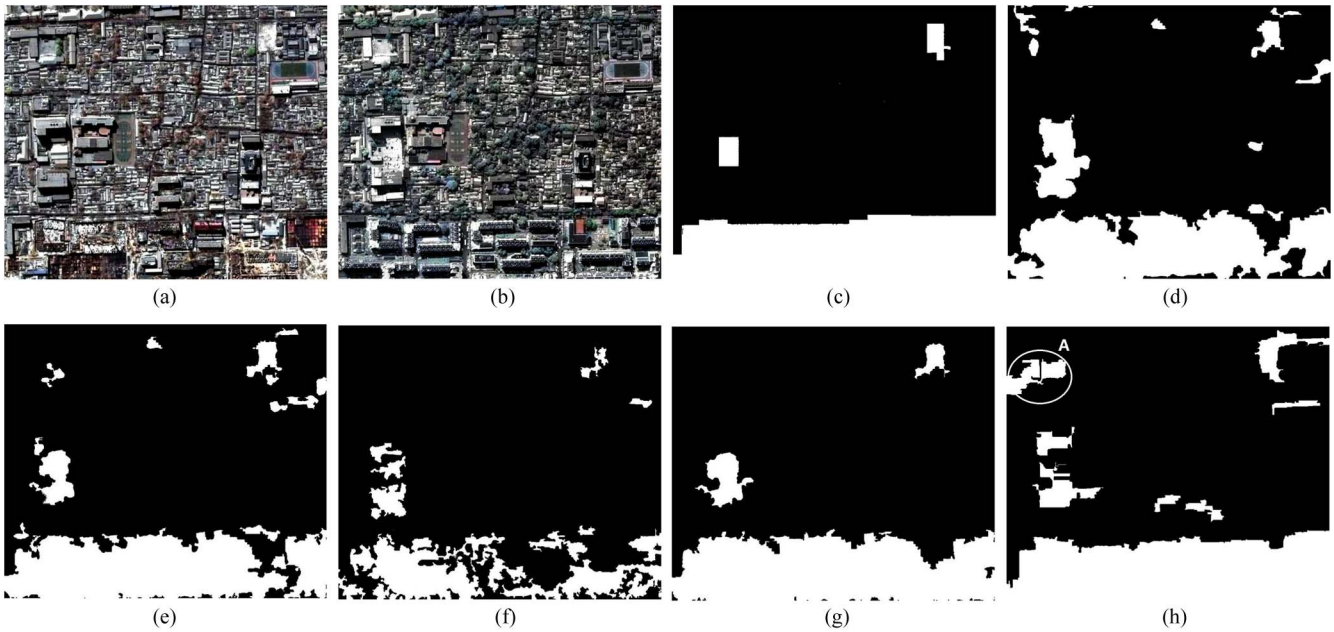


Fig. 3. Performance comparison: dataset 1. (a) and (b): multi-temporal images (882 × 1041 pixels), (c): ground truth, (d): result by DAISY-dif, (e): result by cDAISY-dif, (f): result by (x,y), (g): result by (x,y,cDAISY-dif), (h): result by obj-itsvm.

propagating good training samples and removing bad training samples iteratively. In this letter, iterated transductive SVM [16] is used to classify the change field.

IV. Experiments

To validate the effectiveness of the proposed approach, several experiments were carried out. For space limitation, two datasets are discussed in this letter. The images are taken over Beijing (China) acquired by the QuickBird satellite on February 12, 2004 and October 18, 2005, respectively. The datasets consist of multispectral images (R, G, B, and NIR bands, 2.44 m/pixel in 2004 and 2.52 m/pixel in 2005) and panchromatic images (0.61 m/pixel in 2004 and 0.63 m/pixel in 2005). The original images are encoded by 11 bit. In the experiments, we convert the data to TOA reflectance values and re-encode them by 8 bit. The panchromatic image and multispectral image of the same time are registered by the multi-level SIFT matching [15] and merged based on sparse matrix–vector multiplication [17]. For change detection, the multi-temporal pansharpened images are registered by [15]. The pansharpened RGB images are shown in Fig. 3(a), (b) and Fig. 4(a), (b). The sizes of the dataset are 882 × 1041 pixels and 1430 × 1194 pixels respectively.

The main differences between the proposed approach with the related ones [5], [11] are the change measures. To investigate the effectiveness of the proposed change measure, change field is compared with other three change metrics and another object-level change detection approach: DAISY-dif (the pixel-wise DAISY magnitude difference), cDAISY-dif (the compensated pixel-wise DAISY magnitude difference, i.e., the displacement cost item of change field), (x,y) (the displacement item of change field) and obj-itsvm [11] (object level approach based on iterative transductive SVM classifier). For simplicity, the proposed change field is denoted as (x,y,cDAISY-dif). In this letter, the performances of different approaches are evaluated by the following measures [18]: FP (False Positives), FN (False Negatives), TPR (True Positive Rate), FPR (False Positive Rate) and TA (Total Accuracy).

Fig. 4. Dataset 2 (1430 × 1194 pixels).

It is worth noting that the multi-temporal images are taken at different seasons, and the appearances of the vegetation vary significantly. Fortunately, the Normalized Difference Vegetation Index (NDVI), which is computed based on the R and NIR bands of the multispectral images) is comparatively robust to the seasonal variation. In this letter, the vegetation regions are extracted by thresholding the NDVI images. Since the construction and removal of the buildings are of real interest, we remove the vegetation regions common in the multi-temporal images from the change maps in the following experiments. Noting that DAISY is originally designed for the gray images. For the multispectral images considered in this letter, for efficiency, we transform the pansharpened RGB space to the HSV space and build DAISY descriptor on the Hue. Similar to HueSIFT [19], it is scale-invariant and shift-invariant to light intensity. Of course, other techniques presented in [19] (such as the

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT APPROACHES

dataset	approach	FP	FN	TPR (%)	FPR (%)	TA (%)
1	DAISY-dif	46025	35579	82.7	6.5	91.1
	cDAISY-dif	32293	31580	84.7	4.5	93.0
	(x,y)	19527	104499	49.2	2.7	86.5
	(x,y,cDAISY-dif)	23099	17321	91.6	3.2	95.6
	obj-itsvm	47141	11162	94.6	6.6	93.7
2	DAISY-dif	121186	7814	89.9	7.4	92.4
	cDAISY-dif	119147	5649	92.7	7.3	92.7
	(x,y)	69426	8364	89.2	4.3	95.4
	(x,y,cDAISY-dif)	61032	7681	90.0	3.7	96.0
	obj-itsvm	67411	29850	61.3	4.1	94.3

direction combination of DAISY features from each band) can be used without problem, but it will increase the computation complexities. By experiments, we found that for VHR image change detection, the performance differences can be neglected by the different techniques, since DAISY describes the local texture structure.

The performances of different approaches are listed in Table I. The results obtained by different change features are shown in Fig. 3. The ground truth is shown in Fig. 3(c), which is labeled by the experts after careful validation. From Table I, it can be seen that the performance of change field outperforms the other features. In detail, FP of DAISY-dif is 46025. Many false changes are detected due to the misregistration and view angle variation. By cDAISY-dif, FP is reduced to 32293. After the above impacts are compensated, FP is further reduced to 19527 by (x,y). Nevertheless, the combination of (x,y) and cDAISY-dif is superior to either (x,y) or cDAISY-dif, TPR is improved to 91.6%, and TA is increased to 95.6%. Furthermore, the proposed approach is better than obj-itsvm, the underlying reasons lie in the DAISY feature in capturing the complex structures and the change field in achieving the correspondence. The advantages of change field can also by validated by visually comparing the results shown in Fig. 3. By comparing Fig. 3(d) and (e), it can be observed that many false positives are corrected. Similar to cDAISY-dif, the displacement (x,y) is helpful in compensating the inaccurate correspondence. However, many changes are missed. False negatives and false positives are reduced significantly by change field, since it achieves a good balance between the above two items. Object-level change features in obj-itsvm [11] are discriminative in detecting the complex changes, however, as illustrated by the region A in Fig. 3(h), they are less robust to the complex spectral variation and inaccurate correspondence. On the contrary, DAISY describes the local contextual structure, and change field can separate the false changes from the real changes by taking advantages of the discriminative local features and the robust change measure. The advantages of the proposed approach can also be validated by the second dataset. The above comparisons demonstrate the importance of inaccurate correspondence compensation and the effectiveness of change field in measuring the complex changes.

V. CONCLUSION

Change feature representation is the key factor of VHR image change detection. In this letter, change field is proposed to describe the complex changes between images. Compared to the traditional change features, change field has the meaningful interpretation in projecting the high dimensional feature space into 3-D change feature space, and it is promising in separating the false changes from the real changes. To illustrate the effectiveness of change field, a novel change detection approach is presented. It starts with change field computation based on the discriminative local features and local nearest neighbor search, and followed by change field classification based on the progressive transductive SVM. Experiments demonstrate the advantages of the proposed approach. The novelties of the proposed approach lie in the concept of change field to describe the complex changes. The future developments are mainly related to the utilization of newly proposed local features for change field computation and more advanced classifier for change field classification.

REFERENCES

[1] L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2587–2600, Sep. 2006.
[2] C. Huo, Z. Zhou, H. Lu, C. Pan *et al.*, "Fast object-level change detection for vhr images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 118–122, Jan. 2010.
[3] N. Falco, M. D. Mura, F. Bovolo *et al.*, "Change detection in vhr images based on morphological attribute profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 636–640, May 2013.
[4] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.
[5] N. Bourdis, D. Marraud, and H. Sahbi, "Constrained optical flow for aerial image change detection," in *Proc. IGARSS*, 2011, pp. 4176–4179.
[6] M. N. Klaric, B. C. Claywell, G. J. Scott *et al.*, "Geocdx: An automated change detection and exploitation system for high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2067–2086, Apr. 2013.
[7] A. Sundaresan, P. Varshney, and M. Arora, "Robustness of change detection algorithms in the presence of registration errors," *Photogramm. Eng. Remote Sens.*, vol. 73, no. 4, pp. 375–384, 2007.
[8] R. A. Rensink, "Change detection," *Annu. Rev. Psychol.*, vol. 53, pp. 245–277, 2002.
[9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[10] E. Tola, V. Lepetit, and P. Fua, "A fast local descriptor for dense matching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2008, pp. 1–8.
[11] C. Huo, B. Fan, C. Pan, and Z. Zhou, "Combining local features and progressive support vector machine for urban change detection of vhr images," *AnnalsPRS*, vol. I, no. 7, pp. 221–226, 2012.
[12] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.
[13] S. Winder, G. Hua, and M. Brown, "Picking the best daisy," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2009.
[14] J. Kim, C. Liu, F. Sha, and K. Grauman, "Deformable spatial pyramid matching for fast dense correspondences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2013.
[15] C. Huo, C. Pan, L. Huo, and Z. Zhou, "Multilevel sift matching for large-size vhr image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 171–175, Mar. 2012.
[16] V. Sindhwani and S. S. Keerthi, "Large scale semi-supervised linear svms," in *Proc. ACM SIGIR*, 2006, pp. 477–484.
[17] Y. Wang, H. Yan, C. Pan, and S. Xiang, "Image editing based on sparse matrix–vector multiplication," in *Proc. IEEE Conf. Acoust., Speech, Signal Process.*, 2011, pp. 1317–1320.
[18] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness & correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.
[19] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1582–1596, Sep. 2010.