

A New Approach to Change Detection in Multispectral Images by Means of ERGAS Index

Diego Renza, Estibaliz Martinez, and Agueda Arquero

Abstract—In this letter, we propose a novel method for unsupervised change detection (CD) in multitemporal *Erreur Relative Globale Adimensionnelle de Synthèse* (ERGAS) satellite images by using the relative dimensionless global error in synthesis index locally. In order to obtain the change image, the index is calculated around a pixel neighborhood (3×3 window) processing simultaneously all the spectral bands available. With the objective of finding the binary change masks, six thresholding methods are selected. A comparison between the proposed method and the change vector analysis method is reported. The accuracy CD showed in the experimental results demonstrates the effectiveness of the proposed method.

Index Terms—Local relative dimensionless global error in synthesis (ERGAS), multitemporal optical satellite images, thresholding, unsupervised change detection (CD).

I. INTRODUCTION

REMOTE sensing imagery is an excellent tool for rapid mapping applications. In this context, one of the main uses of remote sensing is the detection of natural or anthropic changes. Change detection (CD) is a technique used to determine the change between two or more time periods of a particular object of study. CD involves the use of multitemporal data sets to discriminate areas of land cover change between dates of imaging [1].

Several analytical approaches differing in complexity, computational intensity, and ease of interpretation have been employed in CD studies. In general, these methods can be categorized into two types: The first type covers those methods detecting binary change/no-change information. This performs CD without additional thematic information and produces a CD map in which changed areas are separated from unchanged ones. Here, the comparison is performed directly on the spectral data, resulting in a difference image which is analyzed to separate insignificant from meaningful changes. This category includes image differencing, image rationing, vegetation index differencing, and principal component analysis (PCA), among others [2]. The second type of CD covers those approaches detecting detailed “from-to” change, such as postclassification comparison, CVA, and hybrid CD methods. The basis of CD is quantifying and labeling changes in the spectral space [3].

Even though the ideal goal of CD is often to extract “from-to” change information, in some cases, we only need simple binary change information. Most of the binary CD studies use an appropriate threshold (threshold values in the tails of the histogram) to identify changed areas from a grayscale image; thresholding can be used to create binary images (change masks) [4]. Moreover, it is important to be able to implement new CD algorithms easily and also to provide accurate CD results. Recent works in CD have shown different trends such as multiresolution analysis [5] by means of a particular wavelet, using object-based image analysis techniques [6], neural net-work model using textural images as auxiliary information [7], modeling CD as a transparency computation problem, and assigning to each pixel a set of soft labels [8], using modulated cosine functions and an extended Kalman filter [9], and analyzing the posterior probability by using CVA [10]. However, discrete methods are attractive because changes are typically defined in terms of land cover classes that are familiar to natural resource managers [11]. Therefore, we had a huge interest in developing efficient automatic CD methodologies for all types of land cover changes, focused on urban changes. The main objective of this study was to propose a new unsupervised approach for change/no CD using a spectral index [*Erreur Relative Globale Adimensionnelle de Synthèse* (ERGAS)], which has been modified for local analysis.

II. SPECTRAL COMPARISON BY ERGAS INDEX

ERGAS was proposed by Wald to estimate the overall spectral quality of fused images. Wald looked for a simple number to describe the overall error of a fused product. The quantity filled three requirements [12]:

- 1) independence of units and according to calibration coefficients and instrument gain; it can be applied to unitless quantities or to radiances;
- 2) independence of the number of spectral bands under consideration;
- 3) independence of image resolutions, which allows comparing results obtained in different cases, with different resolutions [12].

In image fusion [13], [14], ERGAS index is applied to two images with different spatial resolution: a reference image (ground truth) and the new fused image. The ERGAS index between two images is given by

$$ERGAS = 100 \frac{h}{l} \sqrt{\frac{1}{N} \sum_{k=1}^N \left(\frac{RMSE(B_k)}{\mu(k)} \right)^2} \quad (1)$$

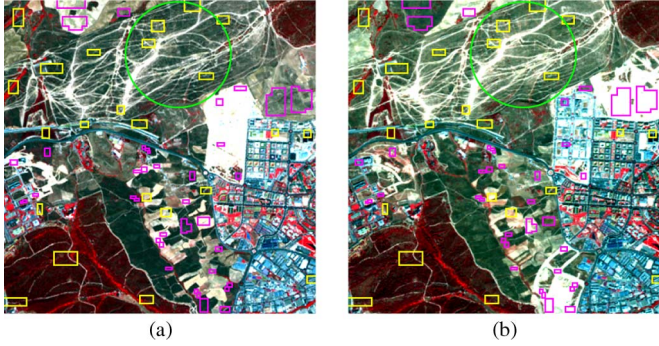


Fig. 1. Data set and test areas: (Magenta) Change and (yellow) no-change areas. False color visualization: NIR, red, and green composite. (a) 2005. (b) 2007.

where h and l denote spatial resolutions (pixel size) of fused and reference images, N denotes the number of spectral bands, k denotes the index for each band, $RMSE(B_k)$ denotes the root-mean-square error (RMSE) for k -band between fused and reference images, and $\mu(k)$ denotes the mean of k -band of reference image. This index is capable of measuring global radiometric distortion between two images. It exhibits a strong tendency to decrease as the quality increases. The closer to zero this number, the better the spectral quality of the final product, so this index provides a quick and accurate insight of the overall spectral similarity between two images.

III. METHODS

A. Study Area

An area with 400×400 pixels of multispectral SPOT5 images (2005 and 2007) were used in this study. We selected anniversary dates to minimize sun angles and seasonal differences. The multitemporal images have a spatial resolution of 10 m. A false color composite of the two multitemporal images is shown in Fig. 1. The upper left corner of the subset is placed at 441280.08E and 4491084.97N (UTM geographic coordinates, zone 30). The study area includes some zones of Community of Madrid (Spain). The scenes mutually exhibit a high proportion of changes due to the urban activity. The land cover types in the study area were defined as urban settlement (buildings and roads), urban planning, Mediterranean forest, vegetated cropland, bare soil, and water (river) according to image interpretation and ground truth. In this letter, we used the visible and near-infrared (NIR) bands of SPOT5.

B. CD Algorithm

In this section, we propose a new CD method based on ERGAS index. In this case, and since the index provides an overall result, it must be modified to work locally, so the resulting mask (CD result) is an image with the same dimensions as the original images. This means that the index is calculated around a neighborhood, which is discussed hereinafter.

1) *Image Normalization*: Due to variations in atmospheric conditions, look/view angles, or sensor parameters that occur between acquisition dates, scenes of the same target area acquired at different times must be normalized to compare

them in an automated fashion [15]. The aim is to minimize the differences between scenes thus enabling more accurate comparisons. A variety of relative radiometric normalization techniques has been developed, such as dark object subtraction and histogram matching, among others. Although SPOT data are supplied with radiometric correction, further processing can be required to get to at-sensor radiance or reflectance, i.e., applying gain and offset [16].

2) *Image Comparison With Local ERGAS*: Single-band CD algorithms are easy to implement and interpret, but their difficulty is to find an algorithm that aggregates the different results for all bands in order to find a single result. We suggest a local ERGAS approach; this approach has the ability to process simultaneously any number of spectral bands to produce a single change image. ERGAS provides a quick and accurate insight of the overall quality of a fused product [12]. However, to obtain a change image, the single quantity given by this index for the whole image does not work. Instead, we propose to use a local method working with neighborhood of each pixel to find a transformation; therefore, we use ERGAS locally.

For local evaluation, the expression given in (1), which means the response at point (x, y) , it is necessary to take into account that RMSE and the mean are the only parameters that range when varying x and y over the entire image. Let $f_k(x, y)$ and $g_k(x, y)$ be the RMSE and mean expressions for each spectral band in each location, so the new equation for local ERGAS is given by

$$ERGAS(x, y) = 100 \sqrt{\frac{1}{N} \sum_{k=1}^N \left(\frac{f_k(x, y)}{g_k(x, y)} \right)^2}. \quad (2)$$

Keep in mind that, in our case, the two images have the same resolution by which the ratio h/l is equal to one. The mean value at each location (x, y) , $g_k(x, y)$ can vary greatly between bright and dark areas of the image. For example, the mean in bright regions is higher compared to that in dark regions, which could cause an ERGAS value greater than the corresponding ERGAS value in a dark region (or vice versa). To minimize this, a first approach can take the mean value for the whole band, such as in (1). However, bright and dark bands can influence the final result as in the previous case. We propose to use the average of all bands

$$g(x, y) = \frac{1}{N} \sum_{k=1}^N \mu(k). \quad (3)$$

Regarding local RMSE, $f_k(x, y)$ creates a new pixel at the coordinates of the neighborhood center; the result is given into a new image. For two images I^1 and I^2 , a window of size $a \times b$, a, b odd, $c = (a - 1)/2$, and $d = (b - 1)/2$, the local RMSE ($f_k(x, y)$) is given by

$$\sqrt{\frac{1}{ab} \sum_{I=-a}^{a_2} \sum_{j=-b}^{b_2} (I_k^1(x+i, y+j) - I_k^2(x+i, y+j))^2}. \quad (4)$$

According to the above, RMSE is the only expression that is calculated in a neighborhood. By working with neighborhoods,

it is necessary to consider the values in the border of the image, where new pixels around the edges of an image are introduced (padding). In (4), for an edge, if you enter the same value for the two images, no matter what values are introduced, this does not affect the RMSE value.

3) *Thresholding*: After the comparison, the grayscale image given by local ERGAS should be thresholded to identify the changed areas. Generally, a specific threshold of the change magnitude is set to determine whether a pixel is changed or not [10]. For selection of thresholds, two methods are often used: to adjust interactively the threshold evaluating the resulting image (trial–error) or by means of automatic strategies with statistical measures, such as selecting a suitable standard deviation from a mean class. The latter has been used in this letter. Consequently, the grayscale image provided by the comparison of ERGAS became (thresholding) a black-and-white (binary) image. We used six different threshold values obtained by applying six thresholding techniques. After that, we had six binary images showing areas of change (white areas) and no change (black areas). Notice that, the higher the pixel value, the more distinct the change. Next, in order to select the best thresholding technique, we analyzed the detection accuracy for changed/no changed regions.

A recent work [17] shows that, for threshold selection, fuzzy ambiguity minimization techniques produce good results, but they need an appropriate window size as an input parameter from the users. For techniques that do not need any user-defined parameters, entropy-based techniques are the best choice. Among these techniques, we choose [4] Entropy-Kapur, Renyi, and Shanbhag thresholding methods. The Entropy-Kapur considers the image foreground and background as two different signal sources, so that, when the sum of the two class entropies reaches its maximum, the image is said to be optimally thresholded. Renyi thresholding is similar to Entropy-Kapur thresholding but using Renyi’s entropy instead. Shanbhag thresholding considers the fuzzy memberships as an indication of how strongly a gray value belongs to the background or to the foreground.

Likewise, we also chose [4] three additional techniques: a clustering-based method, proposed by Otsu, a thresholding based on attribute similarity, proposed by Tsai (Moments), and a fuzzy ambiguity minimization method, proposed by Huang. In Otsu, the optimal threshold is determined by analyzing the behavior of the variances of changed and unchanged classes obtained assuming different threshold values in the range 0 to $L - 1$ (L is the number of gray levels) [17]. In Moments, the gray-level image is considered as the blurred version of an ideal binary image. The thresholding is established so that the first three gray-level moments match the first three moments of the binary image. Huang proposed an index of fuzziness by measuring the distance between the gray-level image and its crisp (binary) version.

4) *Accuracy Assessment*: In an effort to evaluate and compare the proposed algorithm, we try to get some test areas. By comparing these test areas with the results of the proposed method, we can obtain a measure of detection accuracy. Also, these test areas can be compared with the results from other methods, allowing us to contrast results with other methods.

The challenge here is how to establish what exactly the algorithm is expected to produce (ground truth). One option, usually, is by means of an expert human observer; however, multiple expert human observers can differ considerably, even when they are provided with a common set of guidelines [18]. Another option in remote sensing used in classification tasks is to obtain only some test areas, verify its changes by means of field supervision, and then evaluate the results taking into account only those areas. However, when multitemporal images correspond to previous periods, field supervision cannot be performed.

An alternative to get test areas that allow the accuracy assessment can be to apply well-known CD algorithms that result in a change map where changed areas are separated from unchanged ones. In this map, we can select some clearly defined areas for change and not change, which will be the test areas. We chose the latter, and we applied the following procedure: 1) PCA in each multispectral image; 2) absolute difference between the first principal component of the two images; and 3) ISODATA classification (two classes, change/no change) [19]. Using this classified image, we find a good approximation between hit rates and false alarm rates, selecting the test areas shown in Fig. 1, which correspond to field-supervised change and no-change areas. A 5% of test areas for accuracy assessment were collected, both for change and no-change areas. A total of 3986 unchanged pixels (yellow boxes, Fig. 1) and 3998 changed pixels (magenta boxes, Fig. 1) were randomly selected from the interpreted results.

In order to determine the accuracy of the CD approach, the error matrices were constructed, and then, overall accuracy (OA) and kappa index (K) [1], [20] were calculated to assess the whole accuracy of the resulting image. OA is the sum of the correctly classified pixels divided by the total number of test pixels, while K is a statistical measure of the agreement beyond chance, between two maps. In our case, it is between the output map of classification and ground-truth map.

For original images (Fig. 1), our method and CVA were applied, and the same thresholding techniques were evaluated. Recall that our algorithm is proposed to process simultaneously any number of spectral bands. Therefore, it would be appropriate to compare it against a method that also includes several bands. In CVA, a feature vector is generated for each pixel in the image, considering several spectral channels [21]. The modulus of the difference between the two feature vectors at each pixel gives the values of the difference image. CVA determines in n -dimensional spectral space the length and direction of the vector between date 1 and date 2 images and produces an intensity image and a direction image of change. The direction image can be used to classify change. The intensity image must be analyzed (thresholded) to find the change/no-change information.

IV. RESULTS

Histogram matching, dark object subtraction, and conversion to irradiance (through gain and offset) were applied to the input images, in order to minimize radiometric differences. With these corrections, the overall results showed very small

TABLE I
ACCURACY ASSESSMENT FOR ERGAS RESULTS: OA AND K. (TH) THRESHOLD, (C) CHANGE (IN PIXELS), AND (NC) NO CHANGE (IN PIXELS)

TH method		C	NC	Total	OA	K
Huang	C	3998	98	4096	98.77%	0.9754
Th:54	NC	0	3888	3888		
Entropy-kapur	C	3879	0	3879	98.51%	0.9702
Th:97	NC	119	3986	4105		
Moments	C	3993	1	3994	99.92%	0.9985
Th:78	NC	5	3985	3990		
Otsu	C	3993	1	3994	99.92%	0.9985
Th:78	NC	5	3985	3990		
Renyi	C	3939	0	3939	99.26%	0.9852
Th:91	NC	59	3986	4045		
Shanbhag	C	3449	0	3449	93.12%	0.8625
Th:121	NC	549	3986	4535		
	Total	3998	3986	7984		

TABLE II
ACCURACY ASSESSMENT FOR CVA RESULTS: OA AND K. (TH) THRESHOLD, (C) CHANGE (IN PIXELS), AND (NC) NO CHANGE (IN PIXELS)

TH method		C	NC	Total	OA	K
Huang	C	3998	810	4808	89.85%	0.797
Th:67	NC	0	3176	3176		
Entropy-kapur	C	3998	282	4280	96.47%	0.9294
Th:88	NC	0	3704	3704		0.9294
Moments	C	3998	493	4491	93.83%	0.8765
Th:107	NC	0	3493	3493		
Otsu	C	3994	25	4019	99.64%	0.9927
Th:117	NC	4	3961	3965		
Renyi	C	3998	673	4671	91.57%	0.8314
Th:71	NC	0	3313	3313		
Shanbhag	C	3998	915	4913	88.54%	0.7707
Th:64	NC	0	3071	3071		
	Total	3998	3986	7984		

differences, so in this particular case and to reduce the computation time, normalization may be waived, mainly because the two images come from the same sensor, have a similar radiometry, and are captured in anniversary dates.

The local ERGAS and RMSE expressions given in (2) and (4) include sliding-neighborhood operations. These operations require defining the size of the window; square sizes typically are 3×3 , 5×5 , or 9×9 pixels, but other values are acceptable. When applying the local ERGAS, we found that increasing the size of the window decreased the algorithm performance. This may be due to the increase of the size of the window which results in a smoothing effect on the image.

The Multispectral SPOT (XS-SPOT) pixel size in the used images is 10×10 m. By working with neighborhoods, areas would be 900 m^2 (3×3), 2500 m^2 (5×5), and so on. Likewise, by using neighborhoods in CD, the resulting binary image will have a low-pass character which implies that increasing the window size causes the change mask to lose definition. For this reason, practical results with a 3×3 window offer best results. Threshold values in the grayscale image for the two approaches were calculated by means of six algorithms. The results (Tables I and II) and the change masks are shown in Figs. 2 and 3.

By analyzing the results, in accuracy assessment for the CVA-based method, the best one was obtained with Otsu thresholding method, followed by Entropy-Kapur method. For the proposed methodology, the tables show best results for

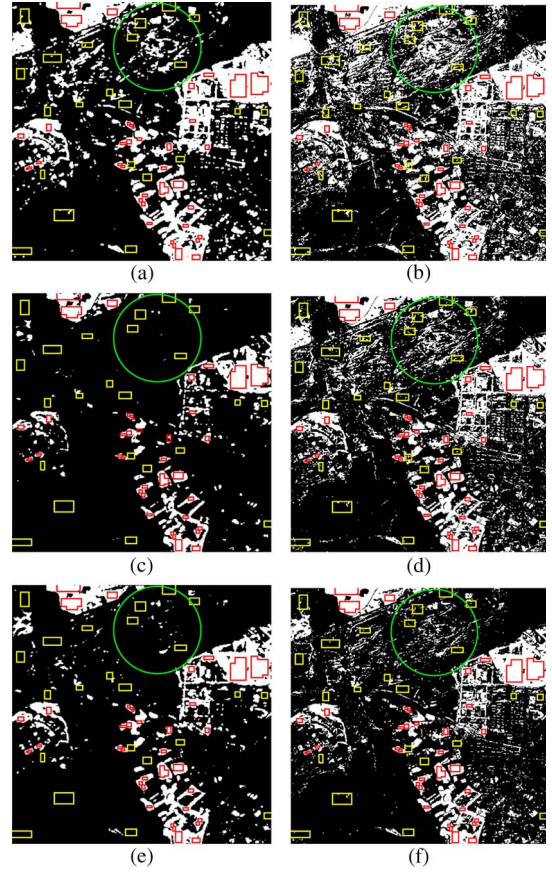


Fig. 2. CD masks obtained through ERGAS-based and CVA techniques corresponding to Huang, Entropy, and Moments thresholding methods. Test areas are displayed in (change) red and (no change) yellow. (a) Huang–our method. (b) Huang–CVA. (c) Entropy–our method. (d) Entropy–CVA. (e) Moments–our method. (f) Moments–CVA.

Moments and Otsu methods. Very close results are obtained for Renyi, Huang, and Entropy-Kapur methods.

Figs. 2 and 3 include the test areas, allowing an initial visual assessment of the CD results. At first glance, one can observe that the masks obtained with the proposed technique show consistent results to actual changes in comparison with the CVA method, even with the same methods of thresholding. For example, visual inspection can be focused on the no changed area marked with the green circle shown in Figs. 1–3. The change masks for our method and CVA-based method with the same threshold can be compared in Figs. 2 and 3. In our method, this area shows lower proportion of falsely detected changed elements compared to the CVA method; this means that a cleaner and easier to interpret change mask was obtained by ERGAS method.

Depending on the thresholding method used, the results may worsen or improve. The higher the threshold, the lower the number of false negatives but the number of false positives may increase. Similarly, the lower the threshold, the smaller the number of false positives but it may increase the number of false negatives. We found that thresholding algorithms that exploit the entropy of the distribution of the gray levels in a scene are good choices for our method. In these algorithms, the maximization of the entropy of the thresholded image is

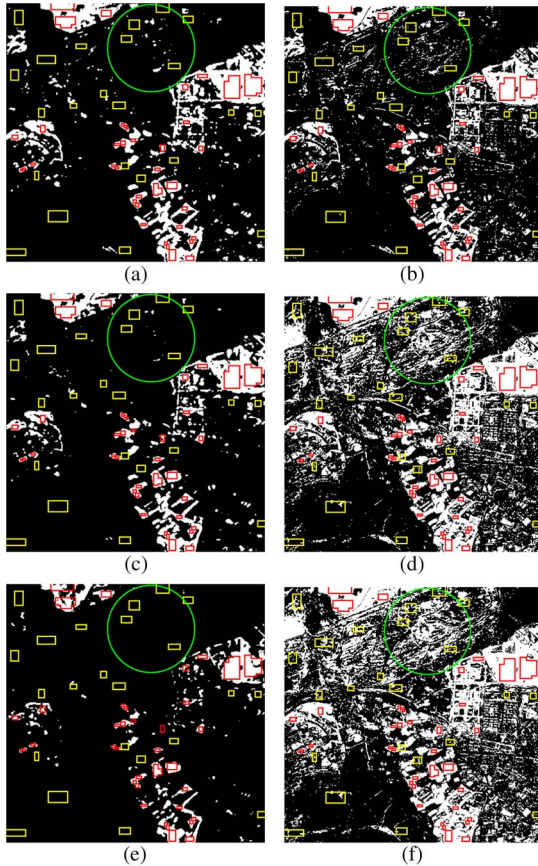


Fig. 3 CD masks obtained through ERGAS-based and CVA techniques corresponding to Otsu, Renyi, and Shanbhag thresholding methods. Test areas are displayed in (change) red and (no change) yellow. (a) Otsu–our method. (b) Otsu–CVA. (c) Renyi–our method. (d) Renyi–CVA. (e) Shanbhag–our method. (f) Shanbhag–CVA.

interpreted as indicative of maximum information transfer. This is the case of entropy-based and Renyi methods.

In the case of the CVA-based method, the choice is not clear, for which we found good results with a clustering-based method (Otsu); however, it has the problem of providing satisfactory results when the numbers of pixels in each class are close to each other. Another thresholding technique that showed good results was Moments.

In both methods, OA and K are good according to the previous results given in Figs. 2 and 3. The ERGAS method gives a slight advantage over the CVA, reducing the detection of false positives; the accuracy of our method has been reported by the OA and K, where they are higher; therefore, the best results were achieved in CD by ERGAS method.

V. CONCLUSION

In this letter, unsupervised CD technique has been developed by thresholding the grayscale change image given by using local ERGAS index, to produce the resultant CD mask. The novel method exploits the benefits of the ERGAS index (the independence of the number of spectral bands) among others. A comparison between the proposed method and the CVA method has been reported. Experimental results have demonstrated the effectiveness of the proposed method. This method has verified

that unsupervised CD in SPOT5 satellite images is possible by using the local ERGAS index; moreover, it does not need any parameters or *a priori* assumptions in finding the difference image. In the future, we will extend our work to the CD on other satellite image data, and we could consider the use of the Spectral Angle Mapper algorithm.

ACKNOWLEDGMENT

The authors would like to thank the three anonymous referees for the valuable comments and suggestions.

REFERENCES

- [1] T. Lillesand, R. Kiefer, and J. Chipman, *Remote Sensing and Image Interpretation*, 2nd ed. Hoboken, NJ: Wiley, 2008.
- [2] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2407, Jun. 2004.
- [3] A. Almutairi and T. Warner, "Change detection accuracy and image properties: A study using simulated data," *Remote Sens.*, vol. 2, no. 6, pp. 1508–1529, Jun. 2010.
- [4] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imag.*, vol. 13, pp. 146–165, Jan. 2004.
- [5] T. Celik and K. Ma, "Unsupervised change detection for satellite images using dual-tree complex wavelet transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1199–1210, Mar. 2010.
- [6] J. Im, J. Jensen, and J. Tullis, "Object-based change detection using correlation image analysis and image segmentation," *Int. J. Remote Sens.*, vol. 29, no. 2, pp. 399–423, Jan. 2008.
- [7] A. Helmy and G. El-Taweel, "Neural network change detection model for satellite images using textural and spectral characteristics," *Amer. J. Eng. Appl. Sci.*, vol. 3, no. 4, pp. 604–610, Apr. 2010.
- [8] W. Luo and H. Li, "Soft-change detection in optical satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 879–883, Sep. 2011.
- [9] W. Kleyhans, J. Olivier, K. Wessels, B. Salmon, F. Van den Bergh, and K. Steenkamp, "Detecting land cover change using an extended Kalman filter on MODIS NDVI time-series data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 507–511, May 2011.
- [10] J. Chen, X. Chen, and X. Cui, "Change vector analysis in posterior probability space: A new method for land cover change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 317–321, Mar. 2011.
- [11] Y. Wang, B. Mitchell, J. Nugranad-Marzilli, G. Bonyngne, Y. Zhou, and G. Shriver, "Remote sensing of land-cover change and landscape context of the national parks: A case study of the northeast temperate network," *Remote Sens. Environ.*, vol. 113, no. 7, pp. 1453–1461, Jul. 2009.
- [12] L. Wald, *Data Fusion: Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*. Paris, France: Presses des MINES, 2002.
- [13] B. Aiazzi, S. Baronti, and M. Selva, "Image fusion through multiresolution oversampled decompositions," in *Image Fusion: Algorithms and Applications*. New York: Academic, 2008, pp. 27–66.
- [14] J. Clevers and R. Zurita-Milla, "Multisensor and multiresolution image fusion using the linear mixing model," in *Image Fusion: Algorithms and Applications*. New York: Academic, 2008, pp. 67–84.
- [15] H. Gang, "Image fusion, image registration, and radiometric normalization for high resolution image processing," Ph.D. dissertation, Univ. New Brunswick, Saint John, NB, Canada, 2007.
- [16] M. El Hajj, A. Bégué, B. Lafrance, O. Hagolle, G. Dedieu, and M. Rumeau, "Relative radiometric normalization and atmospheric correction of a SPOT 5 time series," *Sensors*, vol. 8, no. 4, pp. 2774–2791, Apr. 2008.
- [17] S. Patra, S. Ghosh, and A. Ghosh, "Histogram thresholding for unsupervised change detection of remote sensing images," *Int. J. Remote Sens.*, vol. 32, no. 21, pp. 6071–6089, Nov. 2011.
- [18] R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.
- [19] J. Tou and R. Gonzalez, *Pattern Recognition Principles*, vol. 7. Reading, MA: Addison-Wesley, 1974.
- [20] R. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton, FL: CRC Press, 2009.
- [21] E. Lambin and A. Strahlers, "Change-vector analysis in multitemporal space: A tool to detect and categorize land-cover change processes using high temporal-resolution satellite data," *Remote Sens. Environ.*, vol. 48, no. 2, pp. 231–244, May 1994.