LAND COVER CHANGE DETECTION USING SALIENCY AND WAVELET TRANSFORMATION

Haopeng Zhang[a, b, *], Zhiguo Jiang[a, b], Yan Cheng[c]

[a] Image Processing Center, School of Astronautics, Beihang University, Beijing, 100191, China - (zhanghaopeng, jiangzg)@buaa.edu.cn
[b] Beijing Key Laboratory of Digital Media, Beijing, 100191, China
[c] 93707 PLA Troops, Zhangjiakou, Hebei, 075000, China - cy36152112@126.com

Commission VII, WG VII/5

KEY WORDS: Change detection, Saliency detection, Local entropy, Wavelet transformation

ABSTRACT:

How to obtain accurate difference map remains an open challenge in change detection. To tackle this problem, we propose a change detection method based on saliency detection and wavelet transformation. We do frequency-tuned saliency detection in initial difference image (IDI) obtained by logarithm ratio to get a salient difference image (SDI). Then, we calculate local entropy of SDI to obtain an entropic salient difference image (ESDI). The final difference image (FDI) is the wavelet fusion of IDI and ESDI, and Otsu thresholding is used to extract difference map from FDI. Experimental results validate the effectiveness and feasibility.

1. INTRODUCTION

Land cover change detection is to identify differences in the state of an object or phenomenon by observing it at different times (Singh, 1989). It is a fundamental problem in multi-temporal remote sensing image analysis and plays an important role in both civil and military applications, such as agricultural survey (Bruzzone and Serpico, 1997), forest monitoring (Hame et al., 1998), natural disaster monitoring (Di Martino et al., 2007), urban change analysis (Ridd and Liu, 1998), military reconnaissance etc. Manual labelling can obtain accurate change annotation by observing, comparing, and analyzing two corresponding multi-temporal remote sensing images. However, such work is inefficient and time consuming, and cannot handle the explosive growth of recent remote sensing data. In the past decades, many automatic change detection methods have been studied, including image differencing, vegetation index differencing (Townshend and Justice, 1995), change vector analysis (Bruzzone and Prieto, 2000), principal component analysis (Chang and Wang, 2006), low-rank representation (Cheng et al., 2014), etc. But the key issue that how to obtain accurate difference map remains an open challenge.

Recently, visual attention model has become a research hot-spot in the area of computer vision and pattern recognition. Such model tries to simulate visual attention system of human by computers, and detects salient regions in an image where human eyes focus. For land cover change detection, the area of changed region is usually less than that of unchanged region, and the changed region can be regarded as the salient area by visual attention model. Some researchers (Li, 2009, Yu, 2013) introduced Itti visual attention model (Itti et al., 1998) to change detection. They computed difference map directly on the saliency map, and weakened the details of changed region.

In this paper, we address the problem of detecting accurate difference map using saliency model. As shown in Figure 1, firstly, we use the traditional logarithm ratio method to construct the initial difference image (IDI). Secondly, we detect salient region in IDI using frequency-tuned saliency detection method, and get a salient difference image (SDI). Then, we calculate local entropy of SDI to an entropic salient difference image (ESDI). Finally, IDI and ESDI are fused by wavelet transformation to get final difference image (FDI), and the difference map can be extracted from FDI by Otsu thresholding (Otsu, 1975). There may be noise in IDI obtained by logarithm ratio, and frequency-tuned saliency detection can suppress noise, enhance the change areas in IDI by the saliency and get SDI with the same spatial resolution as IDI. Moreover, dim change region and details of the edges can be kept in ESDI. Thus, when combining IDI with ESDI by wavelet transformation, we can effectively highlight the change region and simultaneously preserve the details. We performed quantitative experiments on two real multi-temporal datasets acquired by Landsat-7 and Landsat-5, and compared with the state-of-the-arts. Experimental results validate the effectiveness and feasibility of the proposed method.

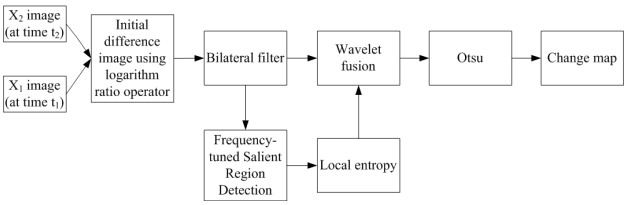


Figure 1: Framework of the proposed method

The rest of the paper is organized as following. In Section 2, we present the proposed change detection method in detail. Experimental results are given in Section 3. Finally, Section 4 comes to a conclusion.

2. METHOD

2.1 Initial Difference Image

Given two registered remote sensing images X_1 and X_2 which separately contains landforms of the same location at different

*Corresponding author

times, the difference image by log-ratio method is as follows

$$D_L = \left|\log \frac{X_2+1}{X_1+1}\right| = |\log(X_2+1) - \log(X_1+1)| \quad (1)$$

We apply a bilateral filter to D_L to reduce noise, and then obtain the initial difference image (IDI) D_I.

2.2 Salient Difference Image

In the initial difference image D_I, the changing area usually takes small ratio of the whole image and with high intensity pixels, while the pixel values in the large unchanging area are close to zero. Thus the changing area can be detected as a salient region by a visual attention model. Following such observation, we introduce frequency-tuned salient region detection (Achanta et al., 2009) to simulate the perception of changing area by human vision, in order to improve change detection result. This method was proposed by Achanta et al. in 2009. It can output full resolution saliency maps with well-defined boundaries of salient objects by retaining substantially more frequency content from the original image.

Give a RGB image I of size $H \times W$, where H and W are the height and width of I, its saliency map S can be calculated by frequency-tuned salient region detection method as the following steps.

Step 1 Apply Gaussian filter to original image I to avoid noise, coding artifacts, and texture patterns. The resulting image after Gaussian filtering is

$$I_g(i,j) = I(i,j) \otimes G \quad (2)$$

where $(i,j), i \in \{1,2,\cdots,H\}, j \in \{1,2,\cdots,W\}$ is the pixel coordinate, $\otimes$ stands for convolution operator, and G is a Gaussian filter in size of 3×3 in this paper.

Step 2 Transfer image I_g from RGB color space to Lab color space, and obtain the luminosity image L and color images a and b.

Step 3 Calculate the mean of image L, a, b respectively as

$$u_L = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} L(i,j) \quad (3)$$

$$u_a = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} a(i,j) \quad (4)$$

$$u_b = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} b(i,j) \quad (5)$$

Step 4 Calculate the square of Euclid distance between the mean of image L, a, b to each pixel in them, i.e. the saliency maps of L, a, b as

$$S_L(i,j) = (L(i,j) - u_L)^2 \quad (6)$$

$$S_a(i,j) = (a(i,j) - u_a)^2 \quad (7)$$

$$S_b(i,j) = (b(i,j) - u_b)^2 \quad (8)$$

Step 5 Final saliency map is the linear combination of the saliency maps of L, a, b as

$$S(i,j) = S_L(i,j) + S_a(i,j) + S_b(i,j) \quad (9)$$

Notice that our initial difference image D_I has only one channel, thus its saliency map, i.e. the salient difference image D_S can be regarded as the saliency map of the luminosity image according to the above method.

2.3 Local Entropy

Changing area in the initial difference image D_I can be enhanced in the salient difference image D_S, but details of edges may become weak. Local entropy can represent the level of intensity variation in a local region of an image. Thus we compute local entropy of D_S to enhance the details in D_I on the premise of retain the salient changing region in D_S, and obtain an entropic salient difference image (ESDI) D_E, of which the value at location (i,j) defined as the entropy value of the 9×9 neighborhood around $D_S(i,j)$ as

$$D_E(i,j) = -\sum_{k=1}^{L} p_{ij}(k) \log p_{ij}(k) \quad (10)$$

$$p_{ij}(k) = r_k(i,j)/81 \quad (11)$$

where L is the number of intensity bins, and $r_k(i,j), k = 1,2,\cdots,L$ is the histogram counts in the 9×9 neighborhood around $D_S(i,j)$.

2.4 Image Fusion Using Wavelet Transformation

Wavelet transformation is a mathematical tool designed for multiresolution signal decomposition (Mallat, 1989) and widely used for image fusion (Chao et al., 2004, Amolins et al., 2007, Hong and Zhang, 2008). In 2D case, let $\{V_l^2, l \in \mathbf{Z}\}$ be a multiresolution analysis of the Hilbert space of functions $L^2(\mathbf{R}^2)$, where $\mathbf{Z}$ and $\mathbf{R}$ denote the set of integers and real numbers respectively. For each $l \in \mathbf{Z}$, the scaling functions $\{\Phi_{l,m_1,m_2} | (m_1,m_2) \in \mathbf{Z}^2\}$ are the orthonormal basis of V_l^2, and the wavelet functions $\{\Psi^{\epsilon}_{l,m_1,m_2} | (m_1,m_2) \in \mathbf{Z}^2, \epsilon = 1,2,3\}$ are the orthonormal basis of $L^2(\mathbf{R}^2)$. Then a 2D image $f(i,j) \in V_l^2$ can be represented by its projection $A_l f(i,j) \in V_l^2$ as

$$\begin{aligned} f(i,j) &= A_l f(i,j) \\ &= A_{l+1} f + D^1_{l+1} f + D^2_{l+1} f + D^3_{l+1} f \end{aligned} \quad (12)$$

where

$$\begin{cases} A_{l+1} f = \sum\limits_{m_1,m_2 \in \mathbf{Z}} C_{l+1,m_1,m_2} \Phi_{l+1,m_1,m_2} \\ D^{\epsilon}_{l+1} f = \sum\limits_{m_1,m_2 \in \mathbf{Z}} D^{\epsilon}_{l+1,m_1,m_2} \Psi^{\epsilon}_{l+1,m_1,m_2}, \epsilon = 1,2,3 \end{cases} \quad (13)$$

Let H_r, G_r and H_c, G_c denote the quadrature mirror filters H, G applying on rows and columns respectively, then wavelet decomposition can be represented as

$$\begin{cases} C_{l+1} = H_r H_c C_l \\ D^1_{l+1} = H_r G_c C_l \\ D^2_{l+1} = G_r H_c C_l \\ D^3_{l+1} = G_r G_c C_l \end{cases} \quad (14)$$

where C_{l+1} is low frequency coefficient and $\{D^{\epsilon}_{l+1}, \epsilon = 1,2,3\}$ are high frequency coefficients. For a remote sensing image X, $C_0(X)$ is X itself. The wavelet reconstruction equation is

$$C_l = H_r^* H_c^* C_{l+1} + H_r^* G_c^* D^1_{l+1} + G_r^* H_c^* D^2_{l+1} + G_r^* G_c^* D^3_{l+1} \quad (15)$$

The key issue for wavelet-based image fusion is the choice of fusion criteria. According to the characteristics of D_I and D_E, we use weighted average criteria, i.e. the low frequency coefficient

of the fused image D_F is

$$C_{l+1,F} = 0.75 \times C_{l+1,I} + 0.25 \times C_{l+1,E} \quad (16)$$

where $C_{l+1,I}$ and $C_{l+1,E}$ are the low frequency coefficients of D_I and D_E respectively, and the high frequency coefficients of D_F are

$$D^{\epsilon}_{l+1,F} = D^{\epsilon}_{l+1,E}, \epsilon = 1, 2, 3 \quad (17)$$

where $\{D^{\epsilon}_{l+1,E}, \epsilon = 1, 2, 3\}$ are the high frequency coefficients of D_E.

In this paper, we practically use Haar wavelet functions and 2-layer wavelet decomposition.

2.5 Difference Map

When we get a fused difference image D_F, Otsu thresholding (Otsu, 1975) is used to obtain the final difference map.

3. EXPERIMENTS

3.1 Datasets

We performed experiments on two real remote sensing datsets to evaluate our method. **Mexico dataset** (Ghosh et al., 2007) is made up of two multi-temporal images acquired by Landsat-7 satellite (channel 4) in an area of Mexico in April 2000 and May 2002, as shown in Figure 2. The two 8-bit images are of size 512×512, and contain 236545 unchanged pixels and 25599 changed pixels in each image. **Sardinia dataset** (Ghosh et al., 2007) is made up of two multi-temporal images acquired by Landsat-5 satellite (channel 4) in September 1995 and July 1996 in the area of Lake Mulargia on the Island of Sardinia (Italy), as shown in Figure 3. The image size is 300×412, and there are 115974 unchanged pixels and 7626 changed pixels.

3.2 Results and Analysis

We compared our method with three state-of-the-art change detection methods, including nonsubsampled contourlet transform (NCT) (Li et al., 2012), dual-tree complex wavelet transform (DTC) (Celik and Ma, 2010), and local entropy saliency model (LESM) (Yu, 2013). Qualitative visual comparison results are shown in Figure 4 and 5, while quantitative comparison results are collected in Table 1 and 2. We can see that results of NCT (Li et al., 2012) contain many false alarm pixels or patches, and DTC (Celik and Ma, 2010) and LESM (Yu, 2013) lose the details of edges. However, our results contain completer changing region with more details and less false alarm pixels. This validates the power of the fusion of entropic salient difference image. Our method achieves the best total change detection error on both datasets, showing its effectiveness.

4. CONCLUSION

In this paper, we have proposed a change detection method based on saliency detection and wavelet transformation. By using wavelet transformation, the initial difference image obtained by logarithm ratio is fused into the entropic salient difference image computed by frequency-tuned saliency detection and local entropy calculation, in order to obtain the final difference image. Then Otsu thresholding is applied to get difference map. Experimental results show that our method can detect accurate changing region with more details and less false alarms.

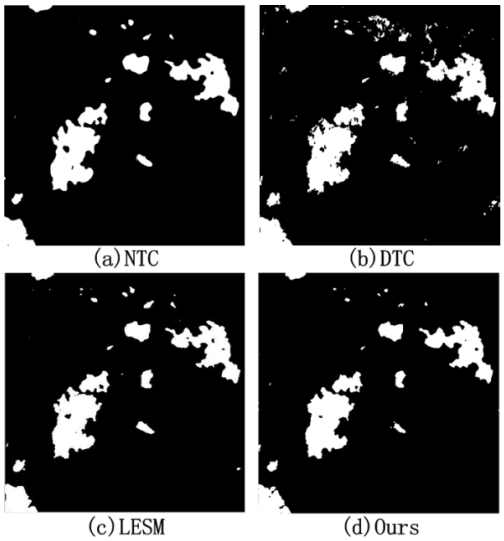


Figure 4: Visual comparison on Mexico dataset

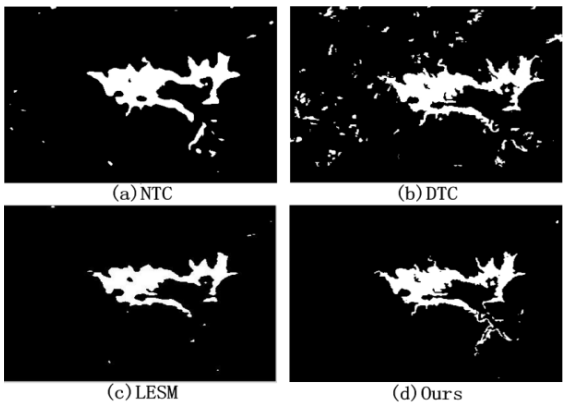


Figure 5: Visual comparison on Sardinia dataset

ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 61501009, 61371134 and 61071137), the Aerospace Science and Technology Innovation Fund of CASC, and the Fundamental Research Funds for the Central Universities.

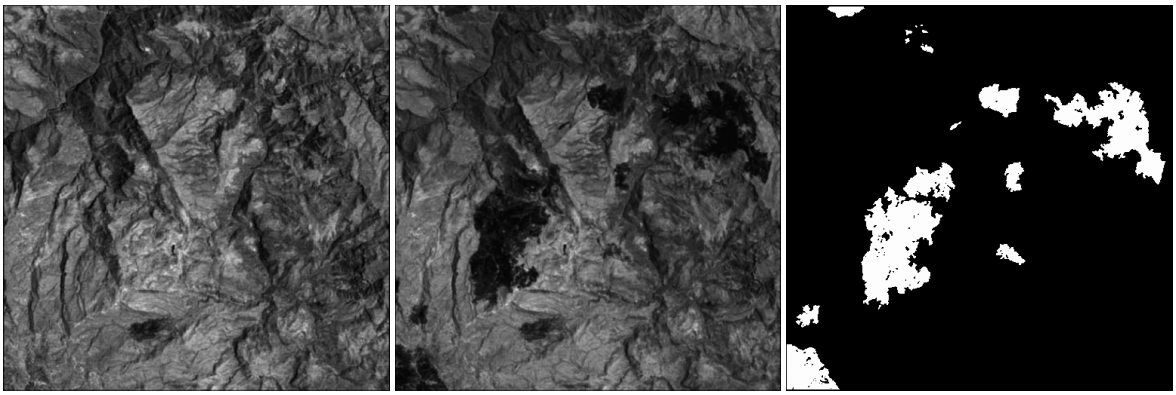
REFERENCES

Achanta, R., Hemami, S., Estrada, F. and Susstrunk, S., 2009. Frequency-tuned salient region detection. In: Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on, IEEE, pp. 1597–1604.

Amolins, K., Zhang, Y. and Dare, P., 2007. Wavelet based image fusion techniques?an introduction, review and comparison. ISPRS Journal of Photogrammetry and Remote Sensing 62(4), pp. 249–263.

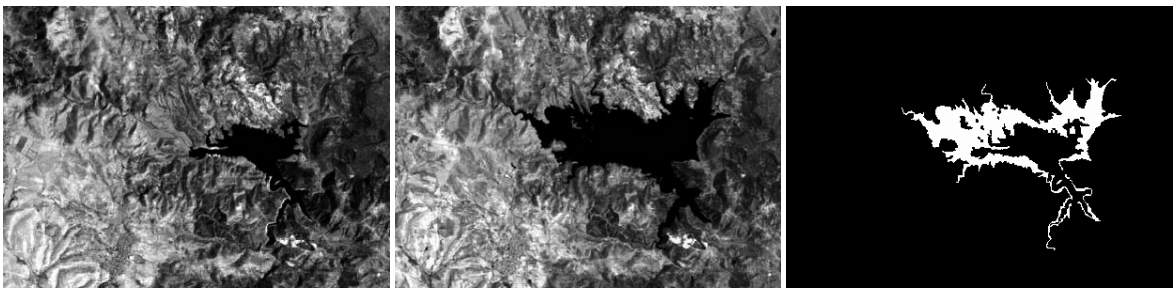
Bruzzone, L. and Prieto, D. F., 2000. Automatic analysis of the difference image for unsupervised change detection. Geoscience and Remote Sensing, IEEE Transactions on 38(3), pp. 1171–1182.

Bruzzone, L. and Serpico, S. B., 1997. An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images. Geoscience and Remote Sensing, IEEE Transactions on 35(4), pp. 858–867.



(a) Image captured on April 2000 (b) Image captured on May 2002 (c) Groud truth of difference map

Figure 2: Mexico dataset



(a) Image captured on September 1995 (b) Image captured on July 1996 (c) Groud truth of difference map

Figure 3: Sardinia dataset

Table 1: Quantitative comparison on Mexico dataset

Methods	False alarm (pixel)	Missed alarm (pixel)	Total error (pixel)
NTC (Li et al., 2012)	2842	1847	4671
DTC (Celik and Ma, 2010)	3698	**834**	4532
LESM (Yu, 2013)	2207	1692	3899
Ours	**1611**	2101	**3712**

Table 2: Quantitative comparison on Sardinia dataset

Methods	False alarm (pixel)	Missed alarm (pixel)	Total error (pixel)
NTC (Li et al., 2012)	3005	583	3588
DTC (Celik and Ma, 2010)	3821	**400**	4221
LESM (Yu, 2013)	**1005**	1210	2215
Ours	1392	662	**2054**

Celik, T. and Ma, K.-K., 2010. Unsupervised change detection for satellite images using dual-tree complex wavelet transform. Geoscience and Remote Sensing, IEEE Transactions on 48(3), pp. 1199–1210.

Chang, C.-I. and Wang, S., 2006. Constrained band selection for hyperspectral imagery. Geoscience and Remote Sensing, IEEE Transactions on 44(6), pp. 1575–1585.

Chao, R., Zhang, K., Li, Y.-j. et al., 2004. An image fusion algorithm using wavelet transform. ACTA ELECTRONICA SINICA. 32(5), pp. 750–753.

Cheng, Y., Jiang, Z., Shi, J., Zhang, H. and Meng, G., 2014. Remote sensing image change detection based on low-rank representation. In: Advances in Image and Graphics Technologies, Springer, pp. 336–344.

Di Martino, G., Iodice, A., Riccio, D. and Ruello, G., 2007. A novel approach for disaster monitoring: fractal models and tools. Geoscience and Remote Sensing, IEEE Transactions on 45(6), pp. 1559–1570.

Ghosh, S., Bruzzone, L., Patra, S., Bovolo, F. and Ghosh, A., 2007. A context-sensitive technique for unsupervised change detection based on hopfield-type neural networks. Geoscience and Remote Sensing, IEEE Transactions on 45(3), pp. 778–789.

Hame, T., Heiler, I. and San Miguel-Ayanz, J., 1998. An unsupervised change detection and recognition system for forestry. International journal of remote sensing 19(6), pp. 1079–1099.

Hong, G. and Zhang, Y., 2008. Comparison and improvement of wavelet-based image fusion. International Journal of Remote Sensing 29(3), pp. 673–691.

Itti, L., Koch, C. and Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis & Machine Intelligence (11), pp. 1254–1259.

Li, S., Fang, L. and Yin, H., 2012. Multitemporal image change detection using a detail-enhancing approach with nonsubsampled

contourlet transform. Geoscience and Remote Sensing Letters, IEEE 9(5), pp. 836–840.

Li, Z., 2009. Research on visual attention models and applicaiton on imagery processing. PhD thesis, Shanghai Jiao Tong University. Shanghai Jiao Tong University.

Mallat, S. G., 1989. A theory for multiresolution signal decomposition: the wavelet representation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 11(7), pp. 674–693.

Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica 11(285-296), pp. 23–27.

Ridd, M. K. and Liu, J., 1998. A comparison of four algorithms for change detection in an urban environment. Remote sensing of environment 63(2), pp. 95–100.

Singh, A., 1989. Review article digital change detection techniques using remotely-sensed data. International journal of remote sensing 10(6), pp. 989–1003.

Townshend, J. and Justice, C., 1995. Spatial variability of images and the monitoring of changes in the normalized difference vegetation index. International Journal of Remote Sensing 16(12), pp. 2187–2195.

Yu, T., 2013. Change detection method of remote sensing images based on graph and sparse representation. Xidian University.