A Novel Framework for Urban Change Detection Using VHR Satellite Images

Weiming Li⁺, Xiaoming Li^{*}, Yihong Wu⁺, Zhanyi Hu⁺

⁺: National Laboratory of Pattern Recognition, Institute of Automation,
Chinese Academy of Sciences, Beijing 100080, P.R. China

^{*}: Department of Mathematics, Shanxi University, Shanxi 030006, P.R. China
Email: {wmli, xmli, yhwu, huzy} @nlpr.ia.ac.cn

Abstract

We present a novel framework for detecting urban changes from a pair of very high resolution (VHR) satellite images such as those taken by satellite Quickbird-II or IKONOS. Image differences due to variations of imaging conditions such as view angle and illumination are distinguished from significant urban changes in the scene. First, we adopt a new image registration method, which makes several useful geometrical constraints available. Then we find changed line segments over time. After that we match Scale Invariant Feature Transform (SIFT) points and generate corresponding regions in order to exclude changed line segments due to parallax. We perform shadow detection to exclude changed line segments due to shadow change. Finally we group the remaining changed line segments into clusters, among which the significant ones form the changed regions as output. Our experiments with real Quickbird-II images show that the proposed method can well detect significant urban changes.

1. Introduction

VHR satellite image pairs taken for the same city at two different times are very useful for analyzing significant urban changes. At present, change detection with such image pairs is primarily performed by human operators, which is rather time consuming and tedious. We aim at developing algorithms that can automatically extract urban change candidates in order to assist human for better efficiency.

The typical observed image changes could be roughly divided into four categories: Significant Urban Changes (SUC), Structural Noise Changes (SNC), View Angle Variation Changes (VAVC), and Shadow Variation Changes (SVC). SUC are changes resulting from construction or dismantlement of city infrastructures such as buildings, roads,

plazas, and grasslands. SUC usually span certain large regions and cause apparent local changes of image structures and textures. SNC are the small unimportant changes which are mainly caused by vehicles in case of the Quickbird-II images. Variations of view angle make some image features corresponding to the same 3D object appear at different locations in the image pair resulting in VAVC. SVC are image changes due to shadow region variations. VAVC and SVC do not constitute real changes in the scene. It is SUC that is the mainly studied point in urban change detection. Therefore, our objective is to extract from image pairs the SUC while excluding SNC, VAVC and SVC.

To detect image changes, many existing change detection methods perform comparisons between various image components such as image pixels[3],[10], lines[11], blocks[1], blobs[9], segments[6], and semantic object parts[12]. As to urban change detection using VHR images, we find that straight line segments are appropriate image features. Line segments are abundant in the urban scene and are less sensitive to illumination changes and local image intensity variations since they integrate the inherent local structural characteristics of the urban objects. However making comparisons between two sets of line segments is a very difficult problem due to the unreliable low-level line extraction process and the complex content of urban scene, especially when VAVC and SVC are present. In our study, we find that VHR image pairs from urban scene provide some useful constraints, by which the problem of change detection could be greatly facilitated. In this work, we incorporate these constraints in each step of our algorithm. Our algorithm contains the following main steps: (1) Image registration. (2) Extraction of line segments. (3) Finding changed line segments over time excluding SNC. (4) Elimination of changed line segments due to VAVC. (5) Elimination of changed line segments due to SVC. (6) Grouping the remaining line segments. Each of the obtained groups covers an image region and such a region is considered due to SUC.

Different from the 3D site model based algorithms [4], our method makes requirement for neither digital elevation model (DEM) nor 3D site model as prior.

In section 2, our method is elaborated step by step. Section 3 reports our experimental results with Quickbird-II VHR images and some analysis. Section 4 are some concluding remarks.

2. Algorithm description

2.1. Image registration

The goal of image registration in most existing change detection methods [3][10][11] is to align the points corresponding to the same object to the same coordinate frame. For VHR image pairs which reveal 3D details of the scene, this goal can be well obtained only when camera view angles are exactly the same across time. Unfortunately in the case of real world, differences in view angles are nearly always present. Different from the existing approaches [12][5], we propose a novel registration scheme which aims at aligning the corresponding objects to approximately the same image locations while still preserving some useful multi-view geometrical constraints.

Most cities are located in planar areas, thus it is reasonable to assume the urban scene under observation is composed of two parts: the ground plane and the 3D architectures standing on this plane. As to the goal of image registration, we propose to align only the corresponding parts from the ground plane and in this way some geometrical constraints for points above the ground plane could be derived.

There exists a homography for corresponding image points from the ground plane. For Quickbird-II image pairs, our experiments show that such a homography transformation is reduced to a similarity transformation. Notice that most image regions are from the ground plane, we adopt Fourier-Mellin transform to globally estimate this similarity transformation. Then, the parameters are used to register the whole images. The registered images are denoted as $Im(t_1)$ and $Im(t_2)$.

Our registration approach preserves some useful geometrical constraints between $Im(t_1)$ and $Im(t_2)$:

- (1) Land cover patterns lying on the ground plane such as roads, plaza, and grassland are aligned to the same coordinates if they are not changed over time.
- (2) For a point P in the 3D scene that is above the ground plane, let P_1 and P_2 be its projections in $Im(t_1)$ and $Im(t_2)$ respectively. Since the imaging process can be approximated by weak perspective projection, the orientation of the vector connecting P_1 and P_2 is a constant relating only to the two view angles. We call this orientation the Main Orientation.

2.2. Extract line segments

We extract line segments from both images using the method of Burns [2]. We eliminate the lines shorter than a threshold so as to suppress SNC. The line segment sets extracted from $Im(t_1)$ and $Im(t_2)$ are denoted by $Line(t_1) = \{l_i^{(1)}\}$ and $Line(t_2) = \{l_j^{(2)}\}$ respectively.

2.3. Find changed lines

Inspired by the work of Burns [2], we propose to make comparisons between line segments from one image and multi-scale gradient orientation field (GOF) of the other image. Multi-scale GOF for $Im(t_i)$ is a set of images: $\{GOF_{Im(t_i)}(x; \sigma_s)\}$, where $x = [x, y]$, $s \in \{1, 2, \dots, 5\}$, and $\sigma_s = 2^{(s-1)/2}$. $GOF_{Im(t_i)}(x; \sigma_s)$ is obtained by calculating the gradient orientation at each point of $Im(t_i)$ after convolving it with Gaussian function with standard deviation σ_s . Then $l_j^{(1)}$ is regarded changed if the following inequality holds:

$$\frac{\max_s \{ \sum_{x \in \Omega(l_i^{(1)})} MA(GOF_{Im(t_2)}(x; \sigma_s), ori(l_i^{(1)})) \}}{\text{length}(l_i^{(1)})} \leq \text{thresh}$$

where $\Omega(l_i^{(1)})$ is the support of $l_i^{(1)}$, $ori(l_i^{(1)})$ is the orientation perpendicular to $l_i^{(1)}$, $\text{length}(l_i^{(1)})$ is the length of $l_i^{(1)}$, and $MA(a, b)$ is a 0-1 function that thresholds the angle difference between the two angles a and b . $l_i^{(2)}$ is treated similarly. Therefore we obtain two sets of changed lines: $Change(t_1) = \{l_i^{(1)}\}$ and $Change(t_2) = \{l_j^{(2)}\}$. These changed lines may due to USC, VAVC or SVC.

2.4. Eliminate changed lines due to VAVC

In spite of view angle variations, the top-down view guarantees that the unchanged objects appear to be similar in the image pairs. Thus by finding the corresponding image regions of similar image appearance we are hopeful to identify the unchanged objects and eliminate VAVC. To characterize local image features, we use SIFT key points [7], which show outstanding performance in capturing distinctive invariant image features [8]. We use the multi-view geometrical constraints preserved by our registration method to improve the matching accuracy and reliability as follows.

Two sets of SIFT key points from the image pair are extracted, denoted by $SIFT(t_1) = \{p_i^{(1)}\}$ and $SIFT(t_2) = \{p_j^{(2)}\}$. Then, $p_j^{(2)}$ is considered as the corresponding point of $p_i^{(1)}$ if it satisfies the following conditions:

- (1) The SIFT descriptor of $p_j^{(2)}$ has larger correlation value with that of $p_i^{(1)}$ than any other key points in $SIFT(t_2)$.

- (2) For $(p_i^{(1)}, p_j^{(2)})$ from the ground plane, the magnitude of the vector connecting $p_i^{(1)}$ and $p_j^{(2)}$ is small.
- (3) For $(p_i^{(1)}, p_j^{(2)})$ from scene point above the ground plane, the orientation of the vector connecting $p_i^{(1)}$ and $p_j^{(2)}$ is consistent with the Main Orientation.

Each SIFT key point has a scale factor which indicates the size of the image region it characterizes. Using this scale factor, we can determine the corresponding image regions from a pair of matched points. In this way we form the corresponding image regions and if changed lines belong to such regions, they will be assigned to VAVC and thus are removed from $Change(t_1)$ and $Change(t_2)$.

2.5. Eliminate changed lines due to SVC

Shadow causes false change detections in two ways: (1) Line segments from shadow contour are labeled as changed lines when shadow regions change across time. (2) Some line segments in one image have their corresponding ones lie in shadow regions of the other image, thus these lines are labeled as change. We use intensity threshold and mathematical morphology method to extract shadow regions from both images. For each $l_i^{(1)}$ in $Change(t_1)$, remove it from $Change(t_1)$ if it satisfies one of the following two conditions:

- (1) $l_i^{(1)}$ lies on the boundary of shadow regions of $Im(t_1)$.
- (2) The corresponding area of $l_i^{(1)}$ in $Im(t_2)$ is within shadow regions of $Im(t_2)$.

Conduct the same procedure for $Change(t_2)$.

2.6. Group changed lines into SUC

We define four distances between two line segments l_i and l_j as $D_k, k=1...4$, as is illustrated in Fig. 1.

For the remaining changed line segments in each $Change(t_i)$, we group them into different clusters such that within each cluster the maximum distance between any two line segments is smaller than a predefined threshold. We observe that the isolated short changed line segments are mainly induced by noise while a cluster of changed line segments spanning certain image regions are usually SUC. Therefore, the clusters whose total line segment lengths are

Figure 1. Four defined distances from two line segments.

Figure 2. (a) Histogram of the orientations of SIFT pairs, (b) corresponding SIFT key points.

smaller than a threshold are removed. Then, from each resulting cluster, the corresponding minimum rectangular region containing all the line segments of this cluster is outputted as the SUC.

3. Experimental results and discussions

We test our framework with two Quickbird-II images taken in Dec. 2003 and Feb. 2005 for the city of Beirut. As samples provided by DigitalGlobe, the images are re-sampled and compressed into 8-bit JPEG format. Image resolution is 60 cm/pixel and each image covers 3.0 square kilometers. After registration with our method, flip-flop visual inspection shows that land cover patterns which are visible on the ground plane in both images are aligned to the same coordinate with displacement error within one pixel.

In Quickbird images, medium vehicle spans a typical length of 10 pixels, thus when thresholding the extracted line segments by this length, short line segments induced by vehicles will be excluded. Larger threshold suppresses more SNC, but also loses detailed SUC.

After extracting SIFT points we take initial matching using only SIFT descriptor correlation. For each matching pair we get the orientation of the vector connecting them, and by calculating the histogram of these orientations, we obtain the Main Orientation. In Fig. 2(a), the peak of the Main Orientation is very sharp, which empirically verifies our geometrical constraints. The geometrical constraints exclude about one fifth of the matching pairs and the rest are consistent with visual inspections. Fig. 2(b) is part of the image pairs showing the details of matching SIFT points.

Fig. 3 (a) and (b) show some the change detection results. This is a typical scene in urban area, there are various urban architectures such as dense buildings with complex roofs (in up right and bottom right parts of the image), parking lot, grasslands. Both view angle and illumination vary between the two times. There are vehicles on the streets and parking lot. The yellow rectangles indicate changes where lines exist in $Im(t_1)$ but disappear in $Im(t_2)$, while the blue

Figure 3. (a)(b)Change detection results,(c)Manual label,(d)Difference image.

rectangles indicate changes where lines exist in $Im(t_2)$ but disappear in $Im(t_1)$. Fig. 3 (c) shows the changed regions by human inspection. By comparison with these regions, 93% of actually changed regions are correctly detected by our automatic detection algorithm and 3% of intact regions are detected as false alarms. We also show the direct difference image of $Im(t_1)$ and $Im(t_2)$ in Fig. 3(d), in which changes due to SUC and other sources are mixed and difficult to separate from each other. Fig. 4 (a),(b), and (c) show the details of the actual changes extracted by our method, where (a) is a square grassland, (b) is the disappearance of a monument and (c) is the removal of walls around a park. (d) is a false detection which is due to the complex pattern formed by both parallax and shadow.

4. Conclusions

We propose a novel framework that can extract significant urban changes from a pair of satellite images while exclude changes due to different imaging conditions such as view angle and illumination. Compared with existing approaches, our method requires no 3D site models and extracts not only building changes but also other significant urban structure changes such as plazas, roads and grasslands. Our experiments using Quickbird-II images show encouraging results. We plan to incorporate more features such as salient image regions into our framework in future work.

This work was supported by the National High Technology R&D Program of China under the grant No: 2005AA114130.

Figure 4. (a)(b)(c)Details of actual changes,(d)An example of false alarm.

References

- [1] M. Borchani, F. Clopper, V. Atalay, and G. Stamoni. Change detection in aerial images. In *Proc. 1st Canadian Conf. on Computer and Robot Vision*, 2004.
- [2] J. B. Burns, A. R. Hanson, and E. M. Riseman. Extracting straight lines. *IEEE Trans. PAMI*, 8(4):425–455, July 1986.
- [3] M. Carlotto. Detection and analysis of change in remotely sensed imagery with application to wide area surveillance. *IEEE Trans. Image Processing*, 6:189–202, Jan. 1997.
- [4] A. Huertas and R. Nevatia. Detecting changes in aerial views of man-made structures. In *IEEE ICCV*, pages 73–80, 1998.
- [5] Y. Kosugi, M. Sakamoto, M. Fukunishi, W. Lu, T. Doihara, and S. Kakumoto. Urban change detection related to earthquakes using an adaptive nonlinear mapping of high-resolution images. *IEEE Geoscience and Remote Sensing Letters*, 1(3):2004, July 152-156.
- [6] J. Li, S. Qian, and X. Chen. Object-oriented method of land cover change detection approach using high spatial resolution remote sensing data. In *Proc. of IGARSS*, pages 3005–3007, 2003.
- [7] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, Nov. 2004.
- [8] K. Mikołajczyk and C. Schmid. A performance evaluation of local descriptors. In *Proc. of CVPR*, pages 257–263, 2003.
- [9] O. Miller, A. Pikaz, and A. Averbuch. Objects based change detection in a pair of gray-level images. *Pattern Recognition*, 38:1976–1992, Nov. 2005.
- [10] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: a systematic survey. *IEEE Trans. Image Processing*, 14(3), Mar. 2005.
- [11] N. C. Rowe and L. L. Grewe. Change detection for linear features in aerial photographs using edge-finding. *IEEE Trans. Geoscience and Remote Sensing*, 39(7):1608–1612, July 2001.
- [12] S. Watanabe and K. Miyajima. Detecting building changes using epipolar constraint from aerial images taken at different positions. In *Proc. of IEEE ICIP*, pages 201–204, 2001.