Change Detection From Media Sharing Community

Naoki Kito, Xiangmin Zhou, Dong Qin, Yongli Ren, Xiuzhen Zhang, and James Thom

School of Science, RMIT University, Melbourne Australia
naokikito1@gmail.com ,{Xiangmin.zhou, Dong.qin, Yongli.ren, Xiuzhen.zhang, James.Thom}@rmit.edu.au

Abstract. This paper investigates how social images and image change detection techniques can be applied to identify the damages caused by natural disasters for disaster assessment. We propose a framework that takes advantages of near duplicate image detection and robust boundary matching for the change detection in disasters. First we perform the near duplicate detection by local interest point-based matching over image pairs. Then, we propose a novel boundary representation model called *relative position annulus* (RPA), which is robust to boundary rotation, location shift and editing operations. A new RPA matching method is proposed by extending *dynamic time wrapping* (DTW) from time to position annulus. We have done extensive experiments to evaluate the high effectiveness and efficiency of our approach.

Keywords: Change Detection · Social Images · Disaster Assessment

1 Introduction

Before, during and after the natural disasters, information about the damages and the current situations are vital for people to make decisions for their next actions. For example, the Nepal earthquake in 2015 did a huge damage to everything, including buildings, roads, infrastructure, and people, which resulted in 8,019 people died and 17,866 people injured; in a Tokyo earthquake, people are still able to walk back home since the earthquake damage the infrastructure, but not the buildings and paths. A recent study [1] found that people would like to know earthquake size and epicentre. Moreover, knowing these information could prevent the secondary and tertiary disasters. It is also found that in Tokyo, only 67.8% of people managed to get back their house on the day of an earthquake, and the rest 32.2% had to become "*homeless*", among them 2% failed to going back home only because they could not find the safe path. People want the information about earthquakes, but there is always the question "How could the people get the information of earthquakes?".

We focus on the problem of *change detection* in sharing communities. In image processing, *change* is defined as the difference between two pixels or the objects in different images. The *difference* varies in different situations. In this paper, the *difference* is limited to the damages to the roads, the buildings or the infrastructures, which are caused by natural disasters. For instance, when a bridge breaks down during an earthquake, the damage to the bridge will be the *change* in this situation. *Change detection* is a process of finding the damages which have been caused by certain natural disasters such as earthquake and flood. Recently, many researchers had approached to detect the damages caused by natural disasters by using the change detection techniques with the aerial images. However, the aerial images consume longer times to retrieve and harder to get compare to the other images. For another, social media is pervasive, and updates very quickly especially on large events, e.g. natural disasters, by millions of people all the time. Thus, we investigate the problem of change detection from social images, so as to let the public aware of the latest situations of natural disasters on the spot. One of the challenges here is the large-scale of social images, which makes current change detection techniques infeasible if not impossible. One of the limitation of using the large-scale images with current change detection techniques is time cost. As existing techniques detect changes based on pixel level comparison [2] without index support or query optimization or both, the time cost for image comparison is high. The second challenge is the unavailability of some special features, like building shady, used in traditional change detection [3]. In shared communities, most social images do not have shady, thus the shady-based matching cannot be conducted. Forcing the existing techniques on the social images will cause low detection quality. Finally, traditional change detection for aerial images suppose the image pairs are known to the same location points, which is not true in media shared communities.

To address these issues, we propose a framework for change detections in sharing communities. First, we conduct PCA-SIFT based matching, which determines if two images are really referring to the same source. Then, we propose a novel boundary representation called *relative position annulus* (RPA), which is robust to the view point rotation and other global transformations of the same objects in images. RPA-based boundary matching between images is performed by extending DTW from series to annulus, which decides if a change has happened after disaster. The rest of the paper is structured as follows: Section 2 reviews the related work followed by the proposed change detection method; Section 4 includes the experiment evaluation; Section 5 concludes the paper.

2 Related Work

We review the related research, including the image copy detection for same source media [18] and change detection. Image copy detection is done by first extracting the descriptors of key points in each image, and counting the number of matched pairs between two compared ones. Examples on descriptors include SIFT[4], PCA-SIFT[5], SURF[6], GLOH[7], and Eff^2[8] etc. In [4], the SIFT is extracted by scale-space extrema selection, keypoint localization, orientation assignment, and descriptor computation. A 128-d vector is computed for each local point. SIFT is invariant to the image variations. However, the matching over it can be expensive due to its high dimensionality. To improve the efficiency, SIFT variants was proposed [5–8]. PCA-SIFT applies principal component analysis to the normalized gradient patches, and is obtained by projecting the gradient image vector computed for each patch into a 36-d space [5]. In [6], SURF was proposed based on the Hessian matrix to approximate the previous descriptor as a vector of length 64. GLOH extended the SIFT by varying the location grid and using PCA to reduce the size [7]. In [8], Eff^2 was proposed by detecting the key points using Difference of Gaussian over different scales, and extracting the information of 8 orientation buckets over each of 3 grid cells around the point. This generates a 72-d vector for each key point. As PCA-SIFT has the stable performance in all situations and has lowest dimensionality[9], we select it in our image detection. To match the descriptor sets, there are mainly two methods: one-many matching and one to one symmetric matching (OOS). In [4], the similarity of two sets is measured by identifying the nearest neighbor of each descriptor based on L_2 distance, and calculating the number of their matched key point pairs. Using this approach, matches over noise key points can be introduced. In [10], Zhao et al. proposed OOS based on a cosine distance-based partial similarity. One key point in a query can only be matched with a single key point of an image data, so the noise matches can be excluded. Considering the superiority of OOS, we choose it for our descriptor similarity.

Object-based change detection (OBCD) algorithms were proposed for satellites, remote sensing and the SAR images to detect the damage and geographical changes from them. In [11], H.Murakami et al. identified changes by subtracting the digital surface model(DSM) from another DSM. In [3], M.Turker et al. used the watershed segmentation to create the segmented building vectors and calculate the shadow area of the segmented building to detect the damages. In [12], L.Matikainen et al. use object-based GIS model data with the overlap analysis algorithm change detection. L.Gong et al.[13] used VHR Terra SAR-X for finding the changes during earthquake. In [14], J.Tu et al. use the 3D GIS model image to detect the building damages during Beichuan earthquake. The 3D GIS model extracts the vectors of building images, and the height of a building is estimated using the shadow detection. Using the satellite, remote and SAR images with damage detection algorithms can achieve high accuracy.Though these images can cover the urban area, there is an issue of retrieving pre-event data sets. Taking all the imagery data for whole country or land is always hard to achieve. Media sharing communities provide great sources for capturing disasters during crisis. Thus, it is demanded to conduct the single pre- and post-event image detection in social communities for disaster management. However, existing methods use typical features that can not be obtained in images from public uploading.

3 Change Detection Over Large Data Sets

We present the change detection over large-scale social images, including how to identify the image pairs from the same source and the changes in images.

3.1 Image Copy Detection

We deploy a PCA-SIFT-based matching over images in dataset, which will be used to decide if two images refer to the same objects by finding the similarity between objects. The main objects are the buildings, statues and other objects which are not moving. Also, the unrelated images selected by the PCA-SIFT algorithm will be removed. The kept image pairs are passed to change detection stage for assessing if any damages had happened in a disaster. We apply the OOS to the similarity between the local descriptor sets of images [15, 16]. Given two descriptors, the similarity of them is measured by the *Cosine* similarity between these two vectors. For two key points from two images, they are match pair candidate if their similarity is bigger than a threshold value. OOS further check if any one of these two points is the nearest neighbor of the other among all the descriptors of its image. If they are nearest neighbors of each other, they are a real matched pair. The final similarity between two images is calculated by the average similarity of all their matched pairs. To improve the key points matching, we use the LIP-IS index [15] as well. All these ensure that effective and efficient PCA-SIFT-based matching is performed.

3.2 Change Detection

Change detection assesses the damages caused during disasters. Intuitively, two pictures to the same location point contains same buildings or other objects, each is described as a boundary. If there is no damage happened at this place, the object boundaries in two images match. Otherwise, if there exists boundary missing or unmatched from the before image to the after one, the damages could have been caused in the disaster. This section proposes a new image object boundary modelling together with a novel boundary matching, which are robust to image transformations, rotations or editing, for effective change detection.

Model image boundary: Existing works model building shadow area boundaries [3], boundary shapes [12] using the coordinates of each pixel falling on the boundary of objects in an image. However, these boundary modelling incurs low effectiveness of detection for social images due to the possible object changes with respect to the viewpoints, rotations, space shift etc. To address this issue, we propose a robust boundary representation, called *relative position annulus* (RPA), which describe each boundary as an annulus of the difference of neighboring edge lengths. Specifically, we exploit sober edge detector to detect a number of boundaries in an image, since it can detect the emphasising edges while reduce the effect of noise edges. Given a boundary consisting of m vertexes $\{v_1, ..., v_m\}$, we represent the boundary as $\{d(v_1, v_2) - d(v_2, v_3), ..., d(v_{m-2}, v_{m-1}) - d(v_{m-1} - v_m), d(v_{m-1}, v_m) - d(v_m, v_1)\}$, where any element can be the start point while other points are ordered clockwise. As such, the RPA will be robust to object rotation, viewpoint change and space shift in social images.

Matching boundaries: As an image may contain multiple objects, each image is described as a set of relative position annulus of multiple boundaries. To assess the damages in a disaster, we need to do two steps matching: (1) the measure between two RPAs; (2) the measure between two images. To further reduce the influence of small noise objects, we only use the top T biggest boundaries for boundary comparison between two images. Given two RPAs, $\mathcal{Q} :< q_1, ..., q_m >$ and $\mathcal{D} :< v_1, ..., v_n >$, we measure their similarity by extending the DTW [19] for RPAs. We consider Q as a series, and D as a set of n series, where each series in the set takes v_i $(i = 1, ..., n)$ as the start point of the boundary and the remaining ones are ordered clockwise. Denote the series to v_i as $\mathcal{D}_i$, and its elements as $< v^i_1, ...v^i_n >$, where $v^i_j = v_{((i+j-1) \pmod n)}$. Then the similarity between $\mathcal{Q}$ and $\mathcal{D}_i$ is measured by:

$$SRPA_i(\mathcal{Q}, \mathcal{D}_i) = \begin{cases} 0 & m = m_1 - 1 \ or \ n = n_1 - 1 \\ max\{SRPA_i(\mathcal{Q}_{m-1}, \mathcal{D}_{n-1}) + Sim(q_m, v^i_n), & \\ \quad SRPA_i(\mathcal{Q}_m, \mathcal{D}_{n-1}), SRPA_i(\mathcal{D}_{m-1}, \mathcal{D}_n)\} & otherwise \end{cases} \tag{1}$$

where Sim is the similarity between q_m and v^i_n computed based on L_1 distance: $Sim(q_m, v^i_n) = 1/(1 + |q_m - v^i_n|)$. The final boundary distance is the maximal DTW between $\mathcal{Q}$ and $\mathcal{D}_i$

$$SRPA = \max_{i=1}^{n} SRPA_i. \tag{2}$$

4 Experiment

This section examines the effectiveness and efficiency of the proposed method.

4.1 Experimental Setup

We use the dataset collected from Flickr by focusing on images relevant to the *Nepal earthquake*, which is also known as *Gorkha earthquake*, in 2015. 100,000 images are collected, which include images *before* and *after* the earthquake. We label the first set of ground-truth image pairs, denoted G_1, for near-duplicate image pair detection (first component of our system), and the second set of ground-truth image pairs, denoted G_2, for change detection (second component of our system). The ground-truth is manually identified by the authors of this paper via careful comparison of all *before* and *after* images. Specifically, for a *ground-truth* in G_1, both *before* and *after* images have to contain at least one same building, but the corresponding image contents, angles, resolutions, colours and light effects could be different. For a *ground-truth* in G_2, there is at least one damaged building found in *after* image comparing with *before* image. We conduct the effectiveness evaluation on 10,000 after and before the earthquake images from the whole dataset, and the efficiency tests on the whole dataset.

To evaluate the effectiveness of algorithms, we used two metrics in [17], the probability of miss detection and false alarm (*Pmiss* and *Pfa*). Specifically, the *missed detection* means that the algorithm fails to detect the ground-truth, and the *false alarm* means the detection of non-target pairs. *Pmiss* and Pfa are defined as follows: $Pmiss = \frac{number\ of\ missed\ detections}{number\ of\ ground\ truth}$, $Pfa = \frac{number\ of\ false\ alarms}{number\ of\ non-targets}$.

We evaluate the efficiency of our approach in terms of the overall time cost of near duplicate image detection, and that of change detection. We compare the time cost of different change detection approaches. Experiments are conducted on Window 7 with Intel(R) Core(TM) i7-4770S CPU (3.4GHz)and 8GB RAM.

4.2 Experimental evaluation

We first test the effect of parameters in change detection, the boundary similarity threshold τ in SRPA and the boundary number threshold T. We then compare the proposed approach with the shape-based change detection(*SBCD*) [12].

Effect of τ: We test the effect of τ in our SRPA measure by varying it from 0.1 to 1. For each τ, we test the $Pmiss$ and Pfa values by setting the number of boundaries considered in each image T to 20, 30, 40, 50. Fig. 1(a) and (b) show the results. It is observed that the $Pmiss$ increases slowly with the increase of τ upto 0.4, then increases dramatically with the further increase of τ after 0.4. On the other hand, the Pfa drops sharply with τ increasing upto 0.4, then decreases slightly after the further increasing of τ. This is because a bigger τ will force a strict constraint on change determination, thus more changes are missed while less false alarms are introduced. Considering a good balance between $Pmiss$ and Pfa, we set the default value of τ to 0.4.

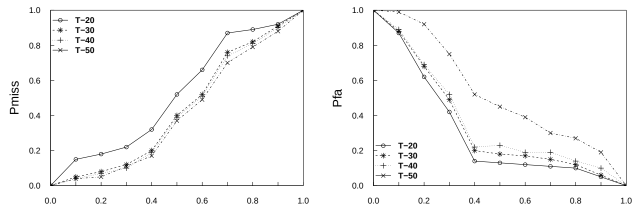

Fig. 1. (a) $Pmiss$ vs. τ (b) Pfa vs. τ

Effect of T: We set τ to it default value and test the effect of T on the effectiveness of change detection by varying T from 5 to 50. Fig. 2 (a) and (b) report the results. Clearly, the $Pmiss$ of the detection drops with T increasing to 25, and keeps steady after that. Meanwhile, its Pfa keeps steady with the increase of T to 25, and increases after 25. To balance the $Pmiss$ and Pfa, we set the default value of T to 25.

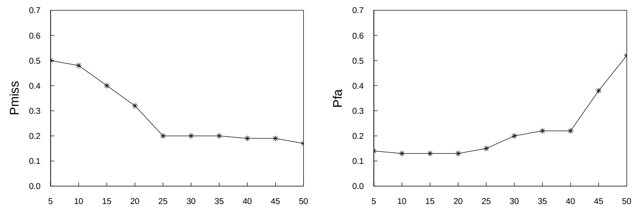

Fig. 2. (a) $Pmiss$ vs. T (b) Pfa vs. T

Table 1. Comparison

	SRPA	SBCD
$Pmiss$	0.14	0.22
Pfa	0.11	0.52

Comparison of change detection techniques: We compare the effectiveness of our SRPA approach with existing shape-based change detection ($SBCD$) [12]. We conduct change detection over the results returned in the first step with the $k = 50$ to avoid missed near duplicate image pairs. We follow all the parameter settings for shape-based matching in [12], and set the parameters τ and T in our SRPA matching to their default values. We test the $Pmiss$ and Pfa values of two approaches. Table 1 reports the results. Clearly our SRPA achieves much lower $Pmiss$ and Pfa, which indicates much better effectiveness, because our RPA is more robust to object variations than the shape representation.

We compare the time cost of our SRPA matching and existing SBCD matching over the near duplicates detected by varying the data size from 10,000 to 100,000. As shown in Fig. 3, our SRPA takes lower time cost than SBCD. This is because SBCD adopts RAP that is a one dimensional annulus. SBCD uses the coordinates of each pixel on the boundary, each is a 2-digital pair, thus more complex than RAP. Compared with SBCD, our SRPA achieves high effectiveness and efficiency, which has proved the superiority of our approach.

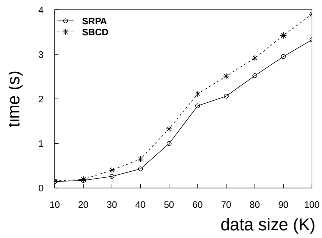

Fig. 3. Comparison

5 Conclusion

In this paper, we study the problem of change detection from media sharing communities for damage assessment in natural disasters. First, we exploit near duplicate detection technique to find each image pair from the same source. Then, we propose a robust boundary representation model together with the matching over it for effective change detection for damage assessment. Finally, we have conducted extensive experiments to evaluate the effectiveness and efficiency of our proposed change detection framework. The experimental results have proved the superiority of our approach over the state-of-the-art method.

References

1. H. U, S. Naoya, N. Ryota, W. Shuntaro, and H. Hidenori, "Questionnaire survey concerning stranded commuters in metropolitan area in the east japan great

earthquake," *Journal of social safety science*, pp. 343–353, nov 2011.
2. M. İlsever and C. Unsalan, *Two-Dimensional Change Detection Methods: Remote Sensing Applications*. SpringerBriefs in Computer Science, Springer London, 2012.
3. M. Turker and E. Sumer, "Building-based damage detection due to earthquake using the watershed segmentation of the post-event aerial images," *Int J Remote Sens*, vol. 29, no. 11, pp. 3073–3089, 2008.
4. D. G. Lowe, "Object recognition from local scale-invariant features," in *ICCV*, vol. 2, pp. 1150–1157, 1999.
5. Y. Ke and R. Sukthankar, "Pca-sift: a more distinctive representation for local image descriptors," in *CVPR* , vol. 2, pp. II–506–II–513, 2004.
6. H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *ECCV*, pp. 404–417, Springer, 2006.
7. K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. on PAMI*, vol. 27, no. 10, pp. 1615–1630, 2005.
8. H. Lejsek, F. H. Ásmundsson, B. T. Jónsson, and L. Amsaleg, "Scalability of local image descriptors: a comparative study," in *MM*, pp. 589–598, 2006.
9. L. Juan and O. Gwun, "A comparison of sift, pca-sift and surf," *IJIP*, vol. 3, no. 4, pp. 143–152, 2009.
10. W.-L. Zhao, C.-W. Ngo, H.-K. Tan, and X. Wu, "Near-duplicate keyframe identification with interest point matching and pattern learning.," *IEEE Trans. on MM*, pp. 1037–1048, 2007.
11. H. Murakami, K. Nakagawa, H. Hasegawa, T. Shibata, and E. Iwanami, "Change detection of buildings using an airborne laser scanner," *ISPRS*, vol. 54, no. 2, pp. 148–152, 1999.
12. L. Matikainen, J. Hyypp, E. Ahokas, L. Markelin, and H. Kaartinen, "Automatic detection of buildings and changes in buildings for updating of maps," *Remote Sensing*, vol. 2, no. 5, pp. 1217–1248, 2010.
13. L. Gong, C. Wang, F. Wu, J. Zhang, H. Zhang, and Q. Li, "Earthquake-induced building damage detection with post-event sub-meter vhr terrasar-x staring spotlight imagery," *Remote Sensing*, vol. 8, no. 11, p. 887, 2016.
14. J. Tu, H. Sui, W. Feng, and Z. Song, "Automatic building damage detection method using high-resolution remote sensing images and 3d gis model.," *ISPRS Annals*, pp. 43–50, 2016.
15. W. Zhao, C. Ngo, H. Tan, and X. Wu, "Near-duplicate keyframe identification with interest point matching and pattern learning," *IEEE Trans. on MM*, vol. 9, no. 5, pp. 1037–1048, 2007.
16. X. Zhou, X. Zhou, L. Chen, A. Bouguettaya, N. Xiao, and J. A. Taylor, "An efficient near-duplicate video shot detection method using shot-based interest points," *IEEE Trans. on MM*, vol. 11, no. 5, pp. 879–891, 2009.
17. X. Zhou, and L. Chen, "Event detection over twitter social media streams," *VLDB Journal*, vol. 23, no. 3, pp. 381-400, 2014.
18. X. Zhou, X. Zhou, L. Chen, Y. Shu, A. Bouguettaya, and J. A. Taylor, "Adaptive subspace symbolization for content-based video detection," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1372–1387, 2010.
19. D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *SIGKDD*, AAAIWS'94, pp. 359–370, AAAI Press, 1994.