Earthquake Geotechnical Engineering for Protection and Development of Environment and Constructions – Silvestri & Moraci (Eds)
 ISBN 978-0-367-14328-2

Extreme natural event damage estimation on buildings and land using image matching and registration

L. Greco, S. Scudero, A. D'Alessandro & G. Vitale
Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Nazionale Terremoti - ONT, Rome, Italy

ABSTRACT: The detection of damage on buildings and land is important in immediate postevent management and in the classification of earthquake or extreme weather event impact. In the first case, it is important to detect the most critical infrastructures affected by the disastrous event to manage the rescue operations. In the second, images are useful to estimate the effects of the extreme natural event and to proper classify the damage. Several works use aerial and satellite images, detecting missing elements and comparing pre-event and after-event views of the area. This approach is useful with a total destruction of elements but is less effective for limited damage. We propose a framework for damage estimation based on couples of ground-based images containing full or partial views of targets to survey. In the proposed method, we perform the recognition and the registration of the pair and finally estimate the damage comparing the transformed images.

1 INTRODUCTION

The detection of structural damage in buildings is crucial in disaster recovery and in preliminary assessments in emergencies, like the immediate post-earthquake. A fast and collaborative assessment of the damage status, in particular the recognition of the most damaged structures and the direct and indirect effects of the earthquake on land, such as induced landslides and rock falls, can help in organizing rescue operations. However, the damage assessment remains a significant challenge because it requires long time, especially in complex urban environments (Dong 2013).

The automatic recognition of structural damage involves several well-known tasks of image processing and computer vision. Most of the applications proposed in the literature exploit multi-temporal satellite and aerial images, either in the visible, infrared, or radar ranges (Ogawa 2000, Ozisik 2004, Vu 2005, Brunner 2010, Rezaeian 2010, Dell'Acqua 2011). However, in the last decade are increasing the applications using processing of oblique imagery taken from Unmanned Aerial Vehicles (UAVs) (Fernandez Galarreta 2014, Duarte 2017). The UAV also allows the 3D reconstruction of the scene and a more thorough assessment of damage (Cusicanqui 2018, Yamazaki 2017).

Applications with ground-based images were not exploited so far since is not always easy to obtain the pre-and-post event pair of images to compare. With the diffusion of databases of georeferenced images such as Google StreetView it is always easier to get the required information to compute the harm of the structure. In this work we present a new framework for damage recognition based on a couple of images containing the pre- and the post-event appearance of a target building or site.

The method encompasses, first of all, the recognition of a (known) target in an image of a damaged area. The source of the second (or more than one) image can be a UAV managed by the civil protection authority or simply ground-based images taken by people nearby the place (social networks, collaborative groups and so on).

The remainder of the paper is organized as follows: Section 2 describes some methods of the state of the art of image registration, features extraction and damage estimation from aerial and satellite images; Section 3 briefly introduce Scale-Invariant Feature Transform (SIFT) descriptor for keypoints extraction from images; Section 4 explains the motivation of homography estimation for planar surfaces and the application of Random Sample Consensus (RANSAC) algorithm; Section 5 focuses on the available datasets of pre- and post- event pictures and on a framework for automatic and participated gathering of images; Section 6 illustrates each step of the proposed method algorithm; Section 7 shows the results of the method, with two examples for earthquakes and rockslides; Section 8 gives some final conclusions on the paper.

2 RELATED WORKS

This paper uses several well-known techniques of Computer Vision and Image Processing for damage estimation and emergency assessment. The most general problem involved in this work is Image Registration, but it can be seen also like image mosaicing, panorama generation from images, or more generally like a geometrical projection.

2.1 *Image registration and features extraction*

Many authors contributed in the field of image registration techniques. An early survey on this task is present in (Brown 1992) and it defines many typical aspects on this task. A more advanced overview is the work of (Zitová 2003) that covers a large number of newer techniques and focuses on feature detection and feature matching.

Feature extraction is the core step for all feature-based image processing algorithms. A funda-mental method to extract interesting points was early proposed by Harris and Stephens (Harris 1988). The method uses the eigenvalues of a squared subset of the image to determine if the point corresponds to an edge, a corner or to a flat region. It is shown (Schmid 2000) that the detector is quite invariant to rotation, scale, illumination changes, and noise.

In this work we use SIFT descriptor (Lowe 2004) to perform image matching to find shared subjects (i.e. buildings) among pairs of images. This descriptor was used for different tasks like recognition (Lowe 1999) and panorama image creation (Brown 2003, Brown 2007).

Other descriptors are present in the state of the art like Maximally Stable Extremal Regions (MSER) (Matas 2004), Speeded Up Robust Feature (SURF) (Bay 2008), Oriented Fast and Rotated Brief (ORB) (Rublee 2011), Principal Components Analysis based SIFT (PCA-SIFT) (Ke 2004). One of the future work of this paper is the implementation and evaluation of these methods in seismic damage recognition and in post-event images generally.

2.2 *Damage estimation from aerial and satellite views*

The majority of works that tries to estimate damage from images exploits multi-temporal datasets of aerial or satellite views, in the visible, infrared, or radar spectra, to obtain the pre-post- event version of a target area.

In (Manfredi 2014) authors discuss a method to refine feature matches using Quickbird images for testing.

In the work (Brunner 2010) damage estimation is performed using a Very High Resolution (VHR) optical pre-event and a VHR Synthetic Aperture Radar (SAR) post-event image as input. The 3D shapes of rectangular and isolated buildings are firstly extracted from pre-event pictures. The resulting shapes obtained from the post-event versions are compared to the expected ones to detect changes. The entire framework provides also a final classification step with damaged/undamaged classes.

Yamazaki and his colleagues (Yamazaki 2004) visually analyzed the effect of the earthquake in Algeria (May 21, 2003) using high-resolution satellite images. The visual inspection proved that results are similar for high damaged zones but start to be different in less damage ones. Artificial Neural Networks were used in (Piscini 2017) to automatically obtain a building collapse ratio from both optical and SAR satellite images.

3 SIFT DESCRIPTOR

SIFT descriptors are generated in intensity (monochromatic) images by finding interesting local keypoints individuating the maxima and the minima of Difference-of-Gaussian in the scale-space pyramid. SIFT algorithm takes different levels (octaves) of gaussian blur on the input image, and computes the difference between the neighboring octaves. Information about orientation vector is then computed for each keypoint, and for each scale.

The SIFT descriptor is defined as a 128-dimensional vector computed by combining the orien-tation of the keypoints in a histogram. SIFT keypoints are extracted in interesting loca-tions as peaks, edges, corners, corresponding to local maxima and minima in Difference-of-Gaussians functions in the neighboring octaves for the image.

As consequence of this, SIFT keypoints are located especially in textured regions, and few or any SIFT keypoints are extracted in homogeneous regions. The resulting descriptor is fully invar-iant to scale and rotation and is robust for affine projections, noise addition, intensity (i.e. illumi-nation) changes, and variation of viewpoint.

An example of extraction and matching of SIFT keypoints is present in Figure 1 (left-top): the starting points of the lines from the left and right images are a subset the detected interest-ing keypoints, the lines are drawn only for the keypoints that match in the two images. The total number of keypoints extracted from the images is several order of magnitude higher than the matching ones.

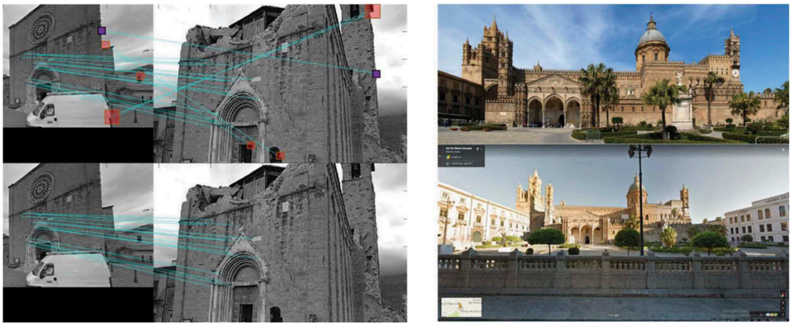

Figure 1. Left: The result of matching points using RANSAC to eliminate outliers (down) and using the original SIFT keypoints (top). Right: Examples of images taken from the web (top) and from Goo-gleMaps (down).

4 RANSAC FOR OUTLIER DETECTION AND HOMOGRAPHY ESTIMATION

The sets of keypoints extracted from the images represent the interesting, or salient, features of the scene. These sets, however, do not perfectly match because the scenes are different (con-tent, point of view and so on).

The first step in image registration using features is to match the extracted sets individuating the shared parts of the two images.

However, it is possible to have wrong matches in those parts of the images that share visu-ally similar features (like corners, edges) but that are not consistent with the geometry of the scene (i.e. the correspondence of the target damaged building).

In this work we deal with the problem of matching scenes to detect damaged (missing, changed) parts as a preliminary step in the automatic classification and detection of damage.

Buildings, in general, are composed by planar surfaces (facades) so, among the possible geo-metrical transformations, we search for a planar projective one that links the targets in the two images.

Such transformation is called homography and can be defined as follows: $[x_1 y_1 1]^T$ and $[x_2 y_2 1]^T$ as two corresponding points (in homogeneous coordinates) of the two images that belong to the same planar (real) surface; and the matrix H as the homographic transformation if the equation

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = H \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \tag{1}$$

$$H = \begin{bmatrix} h_{11}h_{12}h_{13} \\ h_{21}h_{22}h_{23} \\ h_{31}h_{32}h_{33} \end{bmatrix} \tag{2}$$

holds for every pairs of points of the detected surface.

The matrix H can be multiplied by each non-null value, so the real degrees of freedom are 8, and it can be estimated using at least 4 corresponding points.

To determine whether a point should be considered as an outlier (i.e. a wrong match), it is necessary to perform a check on the matched points to exclude the ones that are not consistent with a geometrical projection. A classical method to reject outliers is RANSAC (Fischler 1981) that can be summarized in the following steps:

Ransac Algorithm for Homography

1. From the initial set of candidates, select randomly n points to estimate the parameters for the model;
2. Solve for the estimated parameters, obtaining the homography H;
3. Calculate inliers/outliers according to H;
4. Estimate H using only inliers;
5. Repeat until the Sum of Squared Differences of inliers in the two images is larger than a threshold T.

Figure 1 shows the result of the application of RANSAC algorithm for fitting homography between two images of the same building. The upper part of the left image contains the appearance of the church before the seismic event and the same church after the event. The pair of images matches by few points (19, represented by the lines) but five of these matches are wrong (marked with boxes). In the couple of images in the bottom of the left figure are shown the matches after RANSAC rejection of the outliers, resulting in a correct correspondence between the facade of the building.

Looking at the violet box, it is easy to understand how homography estimation is useful in match rejection: the detected points are very similar (they are placed in an edge of the building with the sky as background) but the match does not fit the transformation of the (majority) of the remaining points, actually belonging to the same facade.

5 AVAILABLE DATASETS

In a post-earthquake scenario, the optimal image pair to evaluate damage in a building is represented by the pre-event and the post-event appearance of the building taken from the same point of view, from the same camera with equal settings, and in the same illumination conditions. Except for the last constraint, it would be simple to obtain such a pair with a fixed (i.e. surveillance) camera pointed on the building, but unfortunately it is not a common configuration.

In real situations, it is easy to get a pre-event view of a building if it is a monument (Figure 1 right), for example searching pictures in the web, but is quite difficult in the general case. We can use tools like Google Maps or similar to get images of residential buildings or schools, normally with a reduced resolution and a fixed point of view.

On the other hand, it is hard to obtain a post-event picture in real time, because after a severe event the core zone is quickly evacuated and restricted. The dataset of pairs used in this paper was constructed using images taken from press websites (before/after pictures of known buildings or places like squares and streets), generally with high variation in time and

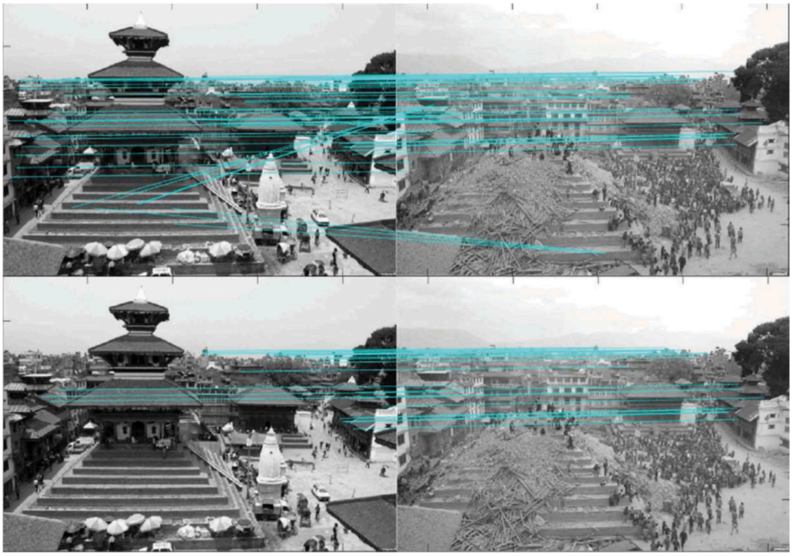

Figure 2. Keypoints matching for a real case using RANSAC (down) and the original matching (top).

viewpoint. The entire framework for damage estimation would involve the part of dataset generation, for example using web-, social network- and voluntary- based retrieval of the images to create correct pre-and-post- event pairs.

6 PROPOSED METHOD

If the pre-event and post-event images are called F and A respectively, the Visual Damage Esti-mation algorithm can be summarized in the following steps:

1. Extract SIFT keypoints from F and A, obtaining *SIFTF* and *SIFTA* sets of points;
2. Find matching points in *SIFT F* and *SIFT A*, obtaining the subsets *M SIFT F* and *M SIFT A* (Figure 3 left);
3. Apply RANSAC algorithm to exclude outliers, obtaining the subsets *R SIFT F* and *R SIFT A*, and to estimate the homography H (Figure 3 right);
4. Calculate the homography of A (called A ’), applying at each pixel position p of the image the transformation $p' = Hp$;
5. Crop the resulting A ’ image and F on the area that the images actually overlap, AC’ = *Crop(A ’)* and FC = *Crop(F)*
6. Calculate the difference image D = |FC-AC’|;
7. For each pixel of D larger than a threshold t, change the value of the corresponding pixel in *AC ’* to a fixed one (red, in our example).

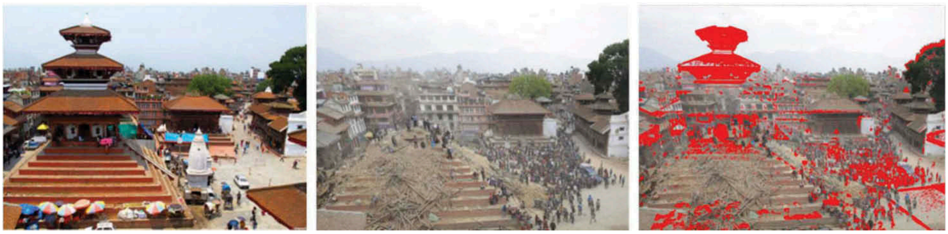

Figure 3. Visual Damage Estimation algorithm result: the difference image (right) shows the detected differences between the original appearance of the building (left) and the post-event image (center).

7 RESULTS

We tested the proposed method for damage assessment on a dataset extracted from the web by hand. The results are consistent with the real damage in the buildings if the image pair actually contains structural elements for feature extraction and registration and if the point of view is enough similar to obtain a consistent homographic projection.The size of the images vary in our dataset, but the average size of the larger dimension is approximately 1000 pixels. In each image pair we search for the match of very distinctive and large part of buildings or rocks as corners and edge discontinuities instead of little details, and as the number of key-points raises with the dimen-sion of the image, it is preferable to use images with a limited resolution (i.e. no more than 2000 pixels for the larger size).

Homography estimation is used for planar correspondence, like for facades, but it can also create a useful projection if the points are far enough from the camera. An example of this behavior is shown in Figure 2. The top part contains the results of SIFT matching of a pair with a missing element and the down part the refined matches using RANSAC. The number of matches decreases and the incorrect ones are discarded.

The result of the registration are present in the left and middle images in Figure 3 and the difference image in the right one using red color to represent the absolute value of the difference.

It is important to notice that the matches are not correspondent to a planar surface and a little error will be introduced in the registration result (noticeable in the top-right tree mismatch red border). Global result, however, is actually reproducing the missing element of the two images, and the red shape of the collapsed pagoda is easy to detect visually.

The difference image is thresholded before visualization (the red points are the ones that exceed two times the mean value of the pixel difference). Changing the threshold creates visually different results introducing noise (decreasing the value) or limiting to strong differences (increasing the value). After several experiments, using the values exceeding two times the mean value seems to be the best one for earthquake scenarios and damage detection.

The application of the damage estimation method on geological events like rockfalls or rock- slides leads to similar results if the elements of interest are quite planar, or if the approximated projection error is bounded. The proposed procedure was applied to several couples of images of the Landslides occurred in Colorado (U.S.A.) in the last decade. Landslides and rockslide in Col-orado are concentrated along the Front Range, central mountains, and western part of the State and are typically associated with areas of significant slope. Landslides and rockslide are natural and ongoing events. Both human activity disrupting the land and periods of significant precipitation, as well as seismic activity, increase the likelihood of fall occurrence. Many landslides and rockslide have not damaged structures, caused deaths or injuries, or affected critical services or facilities. However, some of them have resulted in deaths or significant damage to public infrastructure (Highland 2012).

In the Figure 4 there are two images of a rockslide at Cotopaxi (Colorado). The result of the keypoints matching (in the center) and RANSAC application in the right part of the figure shows showsthat also in this case the method is able to recognize similar points and then to estimate a transformation to (approximately) align the pictures.

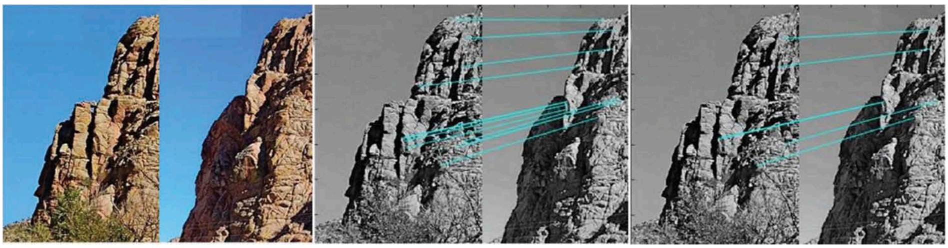

Figure 4. Keypoints extraction for a rockslide: the initial pair (left), the SIFT matching keypoints (center) and the RANSAC refined subset (right).

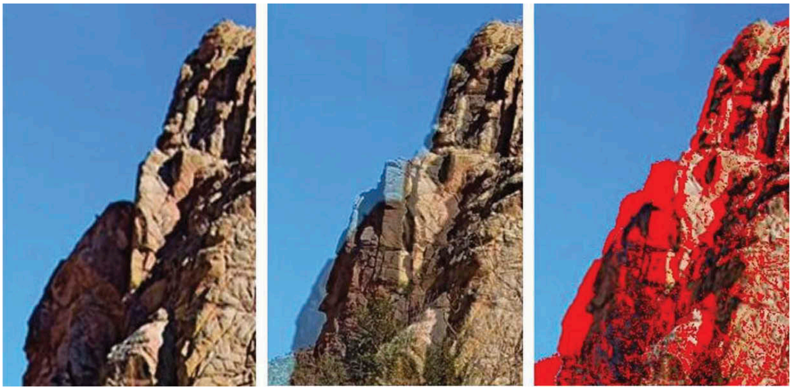

Figure 5. Visual Damage Estimation algorithm result for the rockslide: the superimposed images (center) and the difference image (right).

In the Figure 5 there is the result of the method for the rockslide: the left image is the projection of the after-slide picture, the central one is the superimposition of the original and the projected image. It is possible to notice that, as the input image targets are not planar, a little error on the overlap is present on the border of the rocks, but it is negligible with respect to the collapsed part.

The right image, finally, marks with a large red region the missing part of the rock due to the slide. The remaining smaller red zones are caused by the projection error and illumination changes.

8 CONCLUSIONS

In this work we presented a framework for automatic damage evaluation on building based on the pre-event and post-event images. The problem of damage assessment is crucial in the early hours after a seismic event to organize rescue operations, but is also important for research topics like earthquake intensity distribution.

The proposed method uses feature descriptors referring to single points to identify shared parts among two images (i.e. facades or rocks). The resulting matching points are then divided in outliers/inliers according to a homographic transformation of the target shared parts using RANSAC algorithm.

The use of RANSAC limits our comparison to planar surfaces, but it is a reasonable method for ground-based images, i.e. containing facades, or for pictures taken from similar distances. The results on hand-made couples of images are very encouraging, but the main problem of framework realization is the retrieval of the input pre-event and after-event pair.

To have a good estimation of the error it is necessary to have a large and hand-labeled data-set of image pairs. Based on our knowledge, there are not large datasets publicly available useful to evaluate our method.

As future work we aim to develop an application based on websites, volunteers and social networks for an automatic gathering of such images and to create a larger database of images and to perform better estimations on the error, the precision and the recall of the method.

REFERENCES

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. 2008. Speeded-Up Robust Features (SURF). Computer vision and image understanding 110(3),346-359.

Brown, L.G. 1992. A survey of image registration techniques. ACM Computer Surveys 24(4): 325-376.

Brown, M. & Lowe, D.G. 2003. Recognising Panoramas. IEEE international conference on Computer Vision (ICCV) Vol. 3: 1218.

Brown, M. & Lowe, D.G. 2007. Automatic panoramic image stitching using invariant features. International journal of computer vision 74(1),59-73.
Brunner, D.,Lemoine G.&Bruzzone, L. 2010. Earthquake damage assessment of buildings using VHR optical and SAR imagery. IEEE Transactions on Geoscience and Remote Sensing 48(5): 2403-2420.
Cusicanqui, J., Kerle, N., & Nex, F. 2018. Usability of aerial video footage for 3-D scene reconstruction and structural damage assessment. Natural Hazards and Earth System Sciences 18(6), 1583.
Dell'Acqua, F., Bignami, C., Chini, M., Lisini, G., Polli, D. A. & Stramondo, S. 2011. Earthquake damages rapid mapping by satellite remote sensing data: L'Aquila april 6th, 2009 event. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4(4): 935-943.
Dong, L. & Shan, J. 2013. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. ISPRS Journal of Photogrammetry and Remote Sensing 84: 85-99.
Duarte, D., Nex, F., Kerle, N. & Vosselman, G. 2017. Towards a More Efficient Detection of Earthquake Induced FAÇADE Damages Using Oblique Uav Imagery. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42: 93.
Fernandez Galarreta, J., Kerle, N. & Gerke, M. 2014. UAV-based urban structural damage assessment using object-based image analysis and semantic reasoning. Natural Hazards and Earth System Sciences Discussions, 2: 5603-5645.
Fischler, M.A. &Bolles, R.C. 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6):381–395.
Harris, C. & Stephens, M.J. 1988. A combined corner and edge detector. In Alvey Vision Conference, 15 (50): 147–152.
Highland, L.M. 2012. Landslides in Colorado, USA: Impacts and Loss Estimation for the Year 2010. *USGS Open-File Report* 2012–1204.
Ke, Y. & Sukthankar, R. 2004. PCA-SIFT: A more distinctive representation for local image descriptors. *In* Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004(CVPR 2004) 2: 2. IEEE.
Lowe, D.G. 1999. Object recognition from local scale-invariant features. Proceedings of the seventh IEEE international conference on Computer vision. Vol. 2. IEEE.
Lowe, D.G. 2004. Distinctive image features from scale-invariant keypoints. International journal of computer vision 60(2): 91-110.
Manfredi, M.,Aldrighi, M., &Dell'Acqua, F. 2010.Eigenmethod for Feature Matching of Pre-and Postevent Images Exploiting Adjacency. IEEE Transactions on Geoscience and Remote Sensing 48(7): 2890-2898.
Matas, J., Chum, O., Urban, M., & Pajdla, T.2004. Robust wide-baseline stereo from maximally stable extremal regions. Image and vision computing 22(10): 761-767.
Ogawa, N. & Yamazaki, F.2000. Photo-interpretation of building damage due to earthquakes using aerial photographs, In Proceedings of the 12th world conference on earthquake engineering (No. 1906).
Ozisik, D.& Kerle, N. 2004. Post-earthquake damage assessment using satellite and airborne data in the case of the 1999 Kocaeli earthquake, Turkey. In Proceedings of the XXth ISPRS congress: Geoimagery bridging continents (pp. 686-691).
Piscini, A., Romaniello, V., Bignami, C., & Stramondo, S. 2017. A New Damage Assessment Method by Means of Neural Network and Multi-Sensor Satellite Data. Applied Sciences 7(8): 781.
Rezaeian, M. 2010. Assessment of earthquake damages by image-based techniques. ETH Zurich 107.
Rublee, E.,Rabaud, V.,Konolige, K.&Bradski, G. 2011. ORB: An efficient alternative to SIFT or SURF. IEEE international conference on Computer Vision (ICCV).
Schmid, C., Mohr, R. & Bauckhage, C. 2000. Evaluation of interest point detectors. International Journal of Computer Vision 37(2):151–172.
Vu, T. T., Matsuoka, M.& Yamazaki, F.. 2005. Preliminary results in development of an object-based image analysis method for earthquake damage assessment. In Proceedings of 3rd International workshop Remote Sensing for Post-Disaster Response, Chiba, Japan.
Yamazaki, F., Kouchi, K. I., Kohiyama, M., Muraoka, N., & Matsuoka, M. 2004. Earthquake damage detection using high-resolution satellite images. In Geoscience and Remote Sensing Symposium, IGARSS'04. Proceedings. 2004 IEEE International 4: 2280-2283.
Yamazaki, F., Kubo, K., Tanabe, R., & Liu, W. 2017. Damage assessment and 3d modeling by UAV flights after the 2016 Kumamoto, Japan earthquake. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3182-3185.
Zitová, B. & Flusser, J. 2003. Image registration methods: a survey. Image and Vision Computing 21 (11):977-1000, ISSN 0262-8856.