International Journal of GEOMATE, June 2019, Vol.16, Issue 58, pp.132 - 138
Geotec., Const. Mat. & Env., DOI: https://doi.org/10.21660/2019.58.8218
ISSN: 2186-2982 (Print), 2186-2990 (Online), Japan

3D MODEL-BASED IMAGE REGISTRATION FOR CHANGE DETECTION IN HISTORICAL STRUCTURES VIA UNMANNED AERIAL VEHICLE

Apichat Buatik[1], Luqman Ali[2] and *Krisada Chaiyasarn[3]

[1, 2] Department of Civil Engineering, Thammasat University, Pathumthani, Thailand
[3] Department of Computer and Electrical Engineering, Thammasat University, Pathumthani, Thailand

* Corresponding Author, Received: 21 Oct. 2018, Revised: 15 Dec. 2018, Accepted: 10 Jan. 2019

ABSTRACT: Visual inspection is a common technique to detect and examine the state of health of the structural system. Periodic inspection is carried out to determine if anomalies, such as cracks and surface paint, found in previous visits have changed in appearance over time. The image-based change detection techniques require accurate geometrical and photometrical corrections in pre-processing steps to minimize errors. Although several techniques have been proposed to remove geometrical errors, they still fail to align image correctly, which often results in inaccuracy in a change detection system. In this paper, a change detection system is proposed to tackle this problem. The system acquired images via an unmanned aerial vehicle. Then, the acquired images were manually processed to identify damages, such as cracks, which were used to guide a drone to obtain more images of damage areas for monitoring purpose. The images were then used to obtain a 3D surface model and camera calibration through Structure from Motion (SfM), which were used in the image synthesis technique to obtain an image with identical camera parameters as a queried image for accurate geometrical adjustment. The synthesized images were used to compare with the queried image to see if there were changes between them. In this research project, it was showed that the drone can be used to monitor problematic areas and the image synthetic technique via 3D modeling can be used in geometrical registration to improve a change detection system.

Keywords: Visual Inspection, Change detection, UAVs, Structure from Motion, Historical Buildings.

1. INTRODUCTION

Visual inspection is a common procedure to examine and assess the current state of historical buildings. However, this procedure is laborious and time-consuming as it normally involved inspectors traveling to sites to assess the structures' conditions based on visual appearance. Hence, such a process cannot be conducted frequently due to high labor cost, prone to human-error and site accessibility. Failure to detect problems can lead to disastrous effects.

Many temples in Ayutthaya province, Thailand, have many damages, such as plants growing on structures and surface cracks. These structures may require frequent monitoring and inspection to assess if structures need repair. The change in damages in these structures require closed monitoring to determine if the damages require intervention. This research project aims at improving techniques in change detection to determine if the structures and damages have changed in appearance over time. This project explored the use of an unmanned aerial vehicle for data collection, which is also used to acquire images of problematic areas for monitoring. The acquired images are used in the proposed change detection system, which exploits the technique called Structure from Motion to help to remove geometrical errors, a common problem in change detection systems.

Fig. 1 shows example pictures of cracks in a stupa from Wat Chai Wattanaram obtained from a DSLR camera. This temple has been deteriorating due to natural disasters, in which the Fine Arts department began a renovation program in 1969 and the city was declared a UNESCO World Heritage Site in 1991. Yet, this site still has areas where damages are present, and therefore; frequent inspection and monitoring are required. In this paper, techniques to improve a change detection system are proposed. Firstly images were obtained via an Unmanned Aerial Vehicle (UAV) to build a reference 3D model from images via the Structure from Motion technique. Then, the obtained images were manually identified for interesting areas from the reconstructed 3D model, in which the locations of the areas were used to guide the drone to obtain more images of such areas. The 3D model was then used to create synthetic images, which were used to compare with the real images to create change masks. This process is called geometrical registration, which is essential in removing unwanted noise due to viewpoint difference in a change detection process. The contribution of this paper is, firstly, it was shown that the 3D model from SfM can be used to guide the drone to monitor problematic areas, and secondly, a new technique in geometrical registration is proposed to improve change detection systems.

The rest of the paper is organized as follows, Section 2 is the Literature Review, which summarizes previous work in visual inspection and 3D modeling, and changes detection system. The methodology is explained in Section 3, and Section 4 presents experiments, results, and discussion. The paper ends with the conclusion in Section 5.

Fig.1 Example images of cracks on the surface of stupas visible in many locations around Wat Chai Wattanaram, a temple located in the historic province of Ayutthaya, Thailand.

2. RELATED WORK

2.1 Historical Inspection Based on Image-based 3D Modeling

Three-dimensional digital technologies, such as 3D laser scanners, have been utilized in the conservation works of historical sites, such as in surveying, in archiving and in damage inspection. The technologies offer a number of advantages including speed, convenience, cost-saving, and accuracy. For inspection and damage assessment, Armesto [10] applied a Terrestrial Laser Scanner (TLS) in a masonry bridge to estimate the bridge deformation based on an arch symmetry. Fregonese [12] applied Terrestrial Laser Scanner to monitor out-of-plane displacement of an ancient building by registering two sets of laser scan data to several georeferenced control points and monitor from the difference in displacements. Pieraccini [13] measured the tilt angle of the “Torre del Mangia” (Mangia’s tower) in Siena (Italy) using 3D information from TLS. Costanzo [1] presented a methodology that combined Terrestrial Laser Scanner and infrared thermal image for inspection of St. Augustine Monumental in Calabria (South Italy). Achille [7] demonstrated the applicability of photogrammetry using Unmanned Aerial Vehicle (UAV) to survey the historical structure “Santa Barbara” bell tower in Mantua (Italy). Pesci [3] applied TLS with digital images to detect the trace of restoration in the ancient part of Palazzo d’Accursio in Bologna, Italy. Kouimtzoglou [17] applied to image data from DSLR camera to create a structural plan to restore the Plaka Bridge, an ancient structural heritage in Greece.

The use of digital images in visual inspection is important in providing information about the current state of structural systems. However, its limitations due to inaccessible areas, in which abnormalities may be present, makes the inspection problem hard. This problem may be overcome by the use of an Unmanned Aerial Vehicle (UAV). Achille [7] applied photogrammetry using Unmanned Aerial Vehicle (UAV) to survey the Santa Barbara, which is a historical structure in Mantua (Italy). Bhakapong [4] used a 3D model created from the SfM technique to estimate the tilt angle of Wat Yai Chai Monkol in Ayutthaya Thailand, in which the images were obtained using a drone. It can be seen that UAVs have been applied in the inspection work of historical sites to aid with the data collection process, which can be cumbersome.

2.2 Change Detection

Change detection is a technique that can be used to detect changes in anomalies such as cracks on structures’ surface. A change detection system normally consists of pre-processing steps (i.e. geometrical and photometric adjustments) and a change detection step. The pre-processing steps remove unwanted noise, which is not a real change, from images before the change detection step is applied. Lim [11] proposed a system for monitoring changes in cracks from multi-temporal images. The system is based on a 2D projective transformation that can accurately extract the size of the cracks, which are then monitored in the images as cracks propagate. Delaunoy [8] applies an SFM system to synthesize new views by using a geometric adjustment for change detection in a coral reef. In this work, it is concluded that an SFM system provides an accurate method for synthesizing new views for change detection. Guo [9] proposed the detection of a change (i.e. change detection) is the main component in image interpretation for pipe inspection. Guo [9] presented algorithms for crack detection generally involve a pre-processing step and a crack identification step. The preprocessing step applies image processing techniques to extract potential crack features, such as edges. Chen and Hutchinson, [6] propose a framework for concrete surface crack monitoring and quantification. The method is based on optical flow, which is used to track the movement of cracks. The regions, where the cracks become larger, are labeled as having changed. Chaiyasarn [5] proposed a system for multi-view change detection using images to detect changes of crack on a concrete beam with the use of the SfM for geometrical adjustment. Saur and Krüger [15] present a change detection system using images obtained from a UAV to detect changes between image pairs from video frames. It can be seen that pre-processing steps are important for change detection systems, in which this project is aimed to tackle.

3. METHODOLOGY

Fig. 2 shows the outline of the method proposed in this paper. The system starts with image acquisition (explained in section 3.1), where two sets of images are collected, the first set S_1 taken at time t_1 is used for creating a reference 3D model, and the second set S_2 is taken at the time t_2 is images of interested areas for monitoring purpose. The images from S_1 is input into the *Image-based 3D modelling* module (explained in Section 3.2) to build a 3D model. The images from S_2 are then registered into the 3D model in the *Camera registration* module. In the *Image synthesis* step, a 3D model is rendered using the camera parameters from a queried image I_q, which is a real image from either the set S_1 or S_2 to create a synthesized image I_s (explained Section in 3.3). This process is called geometrical adjustment, which is an essential pre-processing process in change detection to remove noise due to effect from different camera viewpoints. In the *Change detection* module, a change mask is created from I_s and I_q to see if there are changes between images (explained in Section 3.4).

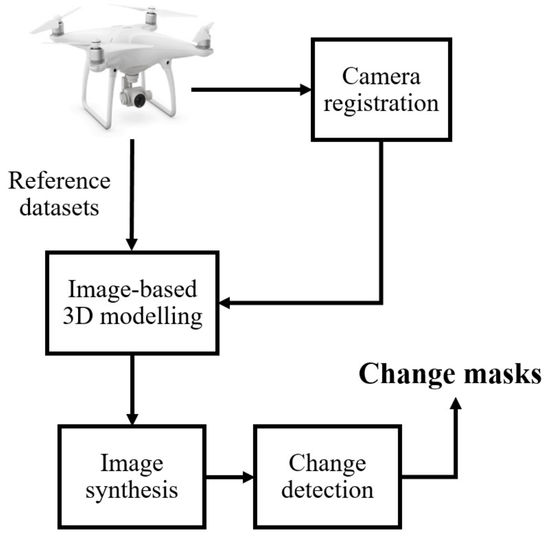


Fig.2 The outline of the proposed system.

3.1 Image Acquisition

Firstly, images are collected to create a reference set S_1 for creating a 3D model. Images were collected by a UAV using the Point of Interest (POI) flight strategy to ensure a full coverage of the stupa and to obtain a highly-detailed 3D model. Fig. 3(a) show the POI strategy, in which the drone flies around an interested object in a circular motion. The flight strategy can be pre-programmed by specifying the radius and height from an interested object. The UAV's camera was programmed to fixate its viewing angle to a stupa. The UAV was programmed to collect images every 2-3 seconds as it moved around the stupa and the images were collected from 3 different heights, i.e. high, mid and ground levels. The ground level was collected manually as the POI cannot be applied as it was too low and the UAV may hit an obstruction on the ground level. Example images obtained from the UAV are shown in Fig. 3(b). In this work, the UAV flight paths were pre-programmed in an iOS application and a DJI Phantom 4 was used. Note that the GPS of each image were also recorded such that the positions images can be registered and estimated in a 3D model.

(a)

(b)

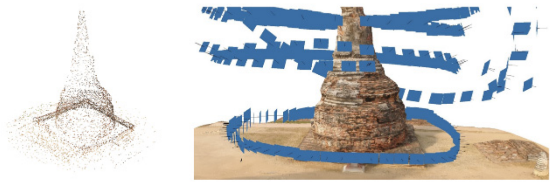
Fig.3 (a) An image showing the Point of Interest flight path for a UAV to collect images around a stupa, (b) Example images of a stupa from Wat Chai Wattanaram.

3.2 Image-based 3D Modeling

The images from Section 3.1 are used to create a 3D model of a stupa in Wat Chai Wattanaram using the Agisoft Photoscan software package, which is based on Structure from Motion and interested readers can refer to Snavely [14] for more detail of the theory and technology. Figure 4(a) shows a sparse point cloud model of a stupa and Fig. 4(b) shows a dense point cloud model with camera calibrations for all images. As shown in Fig. 4(b), each camera also has latitude and longitude information, which helps in creating the 3D model. To obtain a watertight model, a mesh can be created to provide a more realistic 3D model as shown in Fig. 5(a). The mesh is composed of a collection of triangular mesh, in which texture from 2D images is projected onto. The textured model is shown in Fig. 5(b).

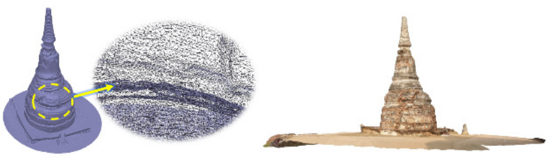
3.2.1 Camera Registration

Images from the reference set S_1 are then identified for interested areas either manually or automatically from the 2D images. An example of an interested area is where cracks appear, in which cracks can be detected automatically from 2D images, but automatic crack detection is beyond the scope of this study. In this work, interested areas were manually identified and the latitude and longitude of the images of the interested areas were used to collect more images from the site as shown in Fig. 6. The new set of images or S_2 are registered onto S_1 by matching similarity between images and by using the GPS information in Agisoft. The 3D model was then re-optimized to update the camera parameters. As shown in Fig. 6, more images were obtained at the interested locations by using the drone to collect more close-up images of the areas.



(a) (b)

Fig.4 (a) a sparse 3D point cloud of a stupa (b) a dense point cloud model with camera calibration.



(a) (b)

Fig.5 (a) 3D mesh model of a stupa, (b) a textured model of a stupa.

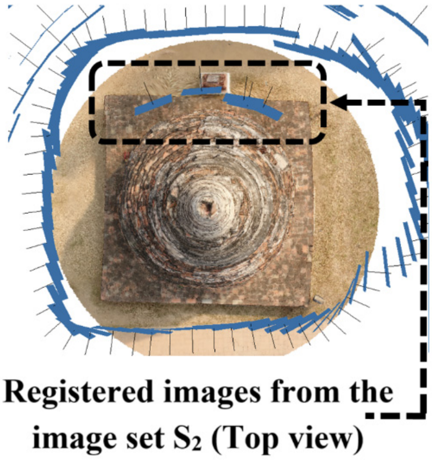


Fig.6 A new set of images is registered onto the reference 3D model, viewed from the top.

3.3 Image Synthesis

A textured 3D model and camera parameters from Section 3.2 are used to synthesize an image from an interested area. This is a geometrical adjustment process, where a real image is compared with a synthesized image as if they are viewed from an identical viewpoint and location. A queried image is a real image, which can be from either S_1 or S_2, called I_q with a rotation matrix R_q and a translation vector t_q and an intrinsic matrix K_q with a set of reference 3D surface as P_r, a synthesized image I_s can be created using the following equation,

$$I_s = K_q[R_q t_q]P_r \quad (1)$$

Fig. 7(a) shows an example of a real queried image I_q and synthesized image I_s, the two images must also be adjusted photometrically. Images were corrected photometrically by the histogram matching technique. Also, a mask was applied to blackout the background on I_q such that the two images are as similar as possible. Fig. 7(b) shows a queried image with lighting adjusted and the background removed. As shown in the Fig. 7(b) and (c), I_q and I_s appear to have identical viewpoints and similar lighting.

(a) (b) (c)

Fig.7 (a) original queried image $\boldsymbol{I_q}$, (b) a queried image $\boldsymbol{I_q}$ with photometrical adjustment and mask applied, and (c) a synthesized image $\boldsymbol{I_s}$.

3.4 Change Detection

3.4.1 Similarity Measure

Similarity measure was applied to compare the similarity between images. The Mean Squared Error (MSE) method is applied, this metric indicates the difference between two images, 0 means perfect similarity and 255 means the two images are dissimilar. MSE can be described as

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j-k(i,j)]^2 \quad (2)$$

Where $I(i,j)$ is a pixel value at i,j of image 1 and $K(i,j)$ is a pixel value of image 2.

3.4.2 Change mask

To create a change mask, pixel values between I_{s} and I_{q} were compared by subtracting gray scale values between the two images using the following equation,

$$D(x) = |I_s(x) - I_q(x)| \quad (3)$$

Then the change mask $B(x)$ is generated according to the following decision rules using the following equation.

$$B(x) = \begin{cases} 1 & if\ |D(x)| > \tau \\ 0 & otherwise \end{cases} \quad (4)$$

τ is a specified threshold. The threshold is currently chosen empirically in order to produce different change masks.

4. EXPERIMENTS AND RESULTS

4.1 Image-based 3D Modelling

In our experiment, images for a reference set $\boldsymbol{S_1}$ were collected using a UAV and the Circular POI flight paths at the 30, 35 and 40 meters was applied. Images were also collected manually on the ground level to ensure the full coverage. The UAV was pre-programmed to take pictures every 2-3 seconds to ensure that an overlap between consecutive images is at least 50%. 520 images of size 5472 x 3642 were collected from a stupa. Also, 5 images for the set $\boldsymbol{S_2}$ were collected, and this set contains closed-up views of images of an interested area.

The reference 3D model of a stupa has 6,658 points for a sparse model, 5,195,908 points for a dense model, and 10,848,033 mesh for a mesh model. Fig. 4 and 5 show the results of our work, the sparse model, the dense model and the textured model. As can be seen from Fig. 4 and 5, the dense and textured model provide good 3D visualization for the temple.

For camera registration, Agisoft can register images using GPS and without using GPS. Table 1 shows the results of registration errors when GPS was applied. As can be seen from the table, the errors from all images from $\boldsymbol{S_2}$ is large and the values are more than 10. This large error is not acceptable. The errors occurs due to that fact that when the dataset $\boldsymbol{S_2}$ was collected, our starting positions was not in a similar location as the starting position when the $\boldsymbol{S_1}$ data was collected. This results in the errors in the GPS information. Therefore, in our experiment, GPS was not used in our registration process. The registration process without GPS works by image matching based on automatically detected keypoints, such as Scale Invariant Feature Transform (SIFT) keypoints (see Snavely [14]) and camera locations are estimated by Triangulation methods. Nevertheless, using the GPS information will provide a better estimate for camera parameters in the real world coordinates.

Table 1 The results of accuracy in image registration using GPS.

Example of dataset	Picture name	Accuracy (m)	Error (m)
$\boldsymbol{S_1}$	DJI_0576	10	1.519775
$\boldsymbol{S_1}$	DJI_0577	10	1.477171
$\boldsymbol{S_2}$	DJI_0646	10	27.562193
$\boldsymbol{S_2}$	DJI_0662	10	27.063962
$\boldsymbol{S_2}$	DJI_0697	10	26.596530

4.2 Change Detection

In this experiment, the change detection technique was applied for queried and synthetic images in 3 different distance from the stupa, far, medium and close. It is worth noting that there are no changes occurred between the datasets $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ as the time difference between the two timeline is not large enough for any change to occur. The change in structures is a slow process, and to observe a change can be difficult. Therefore, in our experiment, no change was expected in our datasets. Table 2 shows the summary results from example images. It can be seen from the table that the far images have the smallest MSE, while the close-up images have the largest MSE. This means that the far images are the most similar. This is due to the fact that the texture of our 3D model is not fine enough to provide sufficient detail in the synthetic images. Our 3D model is created from images that are taken from a far distance, therefore the detail of the 3D model we obtained can only be as fine as the far images can provide. To obtain a more detailed

3D model, more close-up images of the stupa may be required.

Fig. 8 shows the results of change masks between image pairs in the different threshold. The changed pixels are in white and the non-change pixels in black. As mentioned, we expect no change between image pairs. As can be observed from the results, a higher threshold is more accurate as seen in a small number of changed pixels in all image pairs. We can also observe that changed pixels are still present irrespective of threshold setting, and these changed pixels are from noise due to the 3D model that does not provide sufficient details for synthetic images. Nevertheless, we can see that for our result to work, we need to set a higher threshold.

Table 2 A summary table of Mean Square Errors between queried and synthetic images.

Distance from Stupa	Query image	Synthesis image	MSE (pixel^2)
Far (from S_1)			117.54
Medium (from S_1)			250.06
Close-up (from S_2)			2,053.7

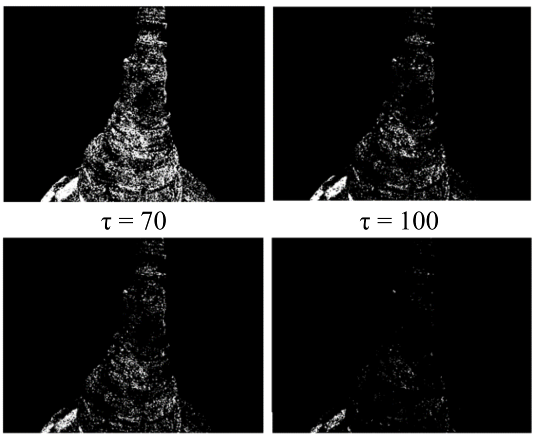


Fig.8 Change masks from sample images (Medium view from Table 2) in various thresholds.

4.3 Discussion

This paper proposed techniques to guide a drone to monitor problem areas and a new technique in geometrical registration for a change detection system. It can be seen that the UAV can be used to monitor problematic areas periodically and obtain flight paths to acquire more images in these areas. The images acquired later can be used to update the 3D model such that the model will be more detailed and allow synthetic images to be more realistic. The system can be combined with damage detection system so that problem areas can be identified automatically, although this is beyond the scope of this paper. In order to get more accurate 3D models, the models should be created using GPS, therefore the UAV must be taken off in similar locations when collecting data at different times. Nevertheless, this research project has demonstrated that we can use a drone to monitor problematic areas.

This research project demonstrates the method to create synthetic images from a 3D model to remove geometrical errors in change detection systems. To improve our results, we require a better 3D model. The current watertight 3D model used in our work may not be sufficient to provide a more realistic synthetic image. Since the watertight model available is created from mesh, in which the texture in each mesh has been smoothened, therefore there will be a loss in the details of a synthetic image due to the smoothing process. To improve the 3D model, more close up images may be required. The proposed change detection algorithm is simple, it may not be suitable for a complex surface, like historical structures. Nevertheless, our method may work well with the simpler surface, such as a concrete surface.

Since changes can be a slow process and to verify a change detection system with real datasets can be cumbersome. Laboratory work is required to verify our method, which we aim to publish in the future.

5. CONCLUSION

The pipeline of a change detection system for historical structures via a guided UAV is presented in this study. The system acquires images via a UAV, in which the acquired images are manually processed to find damages and a 3D model. The UAV can be guided to acquire more image from the problematic areas for periodical inspection and monitoring.

The research proposed method for accurate geometrical adjustment in a change detection system. The adjustment method was achieved by creating a synthetic image from a 3D model. The results of our synthetic image provide a result which can help to remove geometrical errors, although further improvement is required to make a 3D model better to allow a more detail synthetic image for a change detection system to work.

6. ACKNOWLEDGMENTS

The authors would like to thank the Research Division, Faculty of Engineering, Thammasat University, Thailand Research Fund, and the Fine Art Department in Thailand for providing funding and support for the research project.

7. REFERENCES

[1] Costanzo, A., Minasi, M., Casula, G., Musacchio, M. and Buongiorno, M., "Combined Use of Terrestrial Laser Scanning and IR Thermography Applied to a Historical Building", Sensors, 15(1), 2014, 194-213.

[2] Agisoft LLC., Agisoft PhotoScan User Manual : Professional Edition, Version 1.2 downloads: http//:www.agisoft.com/pdf/photoscan-pro_1_2_en.pdf.,2016.

[3] Pesci, A., Bonali, E., Galli, C. and Boschi, E., "Laser scanning and digital imaging for the investigation of an ancient building: Palazzo d'Accursio study case (Bologna, Italy)", Journal of Cultural Heritage, 13(2), 2012, 215 – 220.

[4] Bhadrakom, B. and Chaiyasarn, K., "As-built 3D modeling based on structure from motion from deformation assessment of historical buildings", International Journal of GEOMATE, 11(24), 2016, 2378 – 2384.

[5] Chaiyasarn, K., *Damage Detection and Monitoring for Tunnel Inspection based on Computer Vision*, University of Cambridge, 2014.

[6] Chen, Z. and Hutchinson, T., "Image-based framework for concrete surface crack monitoring and quantification." Advances in Civil Engineering, 2010.

[7] Achille, C., Adami, A., Chiarini, S., Cremonesi, S., Fassi, F., Fregonese, L. and Taffurelli, L., "UAV-Based Photogrammetry and Integrated Technologies for Architectural Applications Methodological Strategies for the After-Quake Survey of Vertical Structures in Mantua (Italy)", Sensors, 15(7), 2015, 15520-15539.

[8] Delaunoy, O., Gracias, N., and Garcia, R., "Towards detecting changes in underwater image sequences." OCEANS 2008-MTS/IEEE Kobe Techno-Ocean, 2008, 1–8.

[9] Guo, W., Soibelman, L., and Garrett, J., "Automated defect detection for sewer pipeline inspection and condition assessment." Automation in Construction, 18(5), 2009, 587–596.

[10] Julia Armesto-González, Belén Riveiro-Rodríguez, Diego González-Aguilera, and M. Teresa RivasBrea., Terrestrial laser scanning intensity data applied to damage detection for historical buildings. Journal of Archaeological Science, 2010, 37(12):3037–3047.

[11] Lim, Y., Kim, G., Yun, K., and Sohn, H., "Monitoring crack changes in concrete structures." Computer-Aided Civil and Infrastructure Engineering, 20, 2005, 52–61.

[12] Luigi Fregonese, Gaia Barbieri, Luigi Biolzi, Massimiliano Bocciarelli, Aronne Frigeri, and Laura Taffurelli. Surveying and monitoring for vulnerability assessment of an ancient building. Sensors, 13(8):9747– 9773, July 2013.

[13] Pieraccini, M., Dei, D., Betti, M., Bartoli, G., Tucci, G. and Guardini, N., "Dynamic identification of historic masonry towers through an expeditious and no-contact approach: Application to the Torre del Mangia in Siena (Italy)", Journal of Cultural HerCombined Use of Terrestrial Laser Scanning and {IR} Thermography Applied to a Historical Buildingitage, 15(3), 2014, 275-282.

[14] Snavely, N., Seitz, S., and Szeliski, R., "Photo tourism: exploring photo collections in 3d." ACM Transactions on Graphics, 25)3), 2006, 835–846.

[15] Saur, G. and Kruger, W., "Change detection in UAV video mosaics combining a feature based approach and extended image differencing". The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016, Volume XLI-B7.

[16] Szeliski, R, Computer vision :algorithms and applications .Springer-Verlag New York, Inc, 2010.

Kouimtzoglou, T., Stathopoulou, E. K., Agrafiotis, P. and Georgopoulos, A., Image-based 3D reconstruction data as an analysis and documentation tool for architects: the case for the Plaka bridge in Greece, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W3, 2017, 1–3 March 2017, Nafplio, Greece.