

# UAV PHOTOGRAMMETRY FOR METRIC EVALUATION OF CONCRETE BRIDGE CRACKS

F. Ioli<sup>1</sup>\*, A. Pinto<sup>1</sup>, L. Pinto<sup>1</sup>

<sup>1</sup> Dept. of Civil and Environmental Engineering (DICA), Politecnico di Milano, Milan, Italy – (francesco.ioli, livio.pinto)@polimi.it, alessandro.pinto@mail.polimi.it

Commission II, WG II/10

**KEY WORDS:** crack detection, UAV, bridge, structural health monitoring, image binarization, concrete structures

## ABSTRACT:

Monitoring cracks opening on concrete bridges is a key aspect for structural health assessment. Digital image processing, combined with Unmanned Aerial Vehicles (UAVs) and photogrammetry, allows for non-contact 3D reconstruction of cracks, reducing costs and potential unsafe factors involved in manual inspections. This paper presents a flexible procedure based on UAV photogrammetry for accurate evaluation of cracks geometry, that can be implemented for periodic structural monitoring. Stereo-pair of images, acquired with UAVs close to the cracked surface, are used to build a scaled photogrammetric model through Structure-from-Motion. Cracks are detected on images by image binarization and digital image processing techniques. Thereafter, one single image is used to reconstruct crack 3D geometry, by back-projecting crack image coordinates on a 3D model of the object. This can be built from the current stereo-pair of images, or based on an existing photogrammetric model, in the case of a periodic monitoring set-up. Crack width is accurately estimated in 3D world. The procedure is tested and evaluated in a case study, obtaining millimetric accurate results, which is in line with the average ground sample distance of the images employed. Results highlight the potentials of UAVs and photogrammetry not only for bridge inspections and damages localization, but also for accurately evaluating cracks geometry and helping structural engineers to assess structure health conditions.

## 1. INTRODUCTION

Many transportation infrastructures, in particular Reinforced Concrete (RC) and Prestressed Concrete (PC) bridges built during the 20th century, are approaching the end of their life cycle, in Italy as well as worldwide (Chaize et al., 2019, ARTBA, 2022). The collapse of the Morandi Bridge in Genova (August 2018, Italy) is just an exemplar case of a wide problem of infrastructure deterioration. Within this framework, monitoring cracks opening in rigid RC or PC structures is a central aspect for periodic structure assessment, as well as it may foretell and, hopefully, prevent disastrous collapses. Traditionally, crack assessment is mainly carried out by visual examination by trained operators. However, inspecting large-scale viaducts piers or beams makes it mandatory to employ under-bridge platforms, often involving high costs and disruption or inconvenience to the infrastructure service. Moreover, visual inspection is usually time-consuming, and it may be subjected to the operator experience.

Digital images have been widely employed for non-contacting crack detection, by replicating human vision. Edge detection is a commonly used method to distinguish cracks boundaries from mainly homogenous background on RGB images (Abdel-Qader et al., 2003, Hutchinson et al., 2006). An alternative approach for crack identification is image binarization, that implies transforming the RGB image into a binary image with one value at crack pixels (Kim et al., 2017). Cracks, in fact, are supposed to be associated to dark pixels, while the background is typically bright. Morphological operations can be then used to enhance crack identification, by exploiting the connectivity of the crack pixels (Tanaka, et al., 1998). An overview of image processing techniques for crack detection was reported by Jahanshahi et al. (2009). The detected crack pixels can be further divided into a skeleton (a chain central pixels representing crack mean direction) and edges.

In recent years, a rising number of research activities has involved the use of Convolutional Neural Networks (CNN) and machine learning techniques to classify images and segment those containing cracks (Cha et al., 2017, Kim et al., 2019, Rao et al., 2020, Li et al., 2020, Kim H. et al., 2022). However, CNN must be properly trained with a conspicuous training set of reference images, and they are most effective when dealing with large dataset of images. Additionally, most of the research was aimed at recognizing the deteriorations on the images, but just few works focused on their metric reconstruction.

The rapid technological advances of Unmanned Aerial Vehicles (UAVs) have opened remarkable opportunities for infrastructure inspections, as they allow images to be taken in proximity to the crack surface, also with large-scale infrastructure. Several example of usage of UAVs for crack investigation and metric reconstruction can be found in recent literature (Rau et al., 2017, Kim et al., 2017, Liu et al., 2019). Moreover, UAVs, combined with modern photogrammetry and Structure-from-Motion (SfM), allowed for building accurate 3D models, mesh models, orthophotos of observed infrastructure (Hackl et al., 2018, Pinto et al., 2020). Recently, Kim et al. (2022) developed a compact and lightweight stereo-system, composed of one wide-angle lens and a telephoto lens, to accurately quantify cracks and with the aim of mounting it onboard UAVs for large-scale infrastructure monitoring.

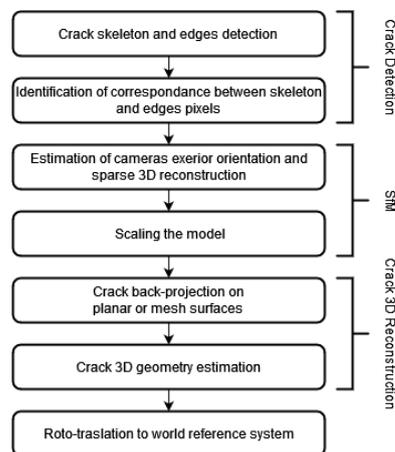
The aim of this paper is presenting a simple and flexible image-based methodology to derive metric information about cracks in RC/PC structures. Images are acquired by UAVs, allowing for inspecting large infrastructure such as highway bridges or viaducts. Crack information is extracted from the images by digital image processing techniques, and crack 3D metric information is derived by a photogrammetric approach. The procedure is validated on a field test and the results assessed by comparing them with independently acquired reference data.

\* Corresponding author

## 2. THE METHOD

The proposed procedure, implemented in MATLAB®, is composed of 3 main parts (Figure 1): crack detection on images; SfM scene reconstruction; crack 3D reconstruction. Stereo-pairs of images are acquired with UAVs close to the cracked surface, and they are employed to build a photogrammetric model by SfM. On images, cracks are detected by image binarization and morphological operations, by exploiting the connectivity of pixels in the cracked regions to identify the skeleton. Crack edges are further detected with edge-detection algorithms. Thereafter, one single image is used for crack metric reconstruction, by back-projecting crack skeleton and edges onto a 3D model of the object. Crack width is estimated in 3D world.

While stereo-pair of images are used to build a 3D model through SfM, one single, oriented, image is used for crack 3D reconstruction, by back-projection pixel information to a 3D model. This is motivated by the fact that crack detection is based on image binarization, and hence, on modelling the crack with a chain of central pixels (i.e., the skeleton) and two edges. Hence, finding homologous skeleton and edges points on binarized images for restituting them by triangulation can be troublesome. The usage of one image for 3D reconstruction overcomes this limit. Moreover, this approach enables the possibility to set up a framework for multi-epochs monitoring of severe cracks. An accurate 3D model of the bridge, in fact, can be built at the first epoch (e.g., by UAV photogrammetry), and then employed to orient subsequent images of the same crack, in order to compute its opening breadth and evaluate the damage evolution in time.

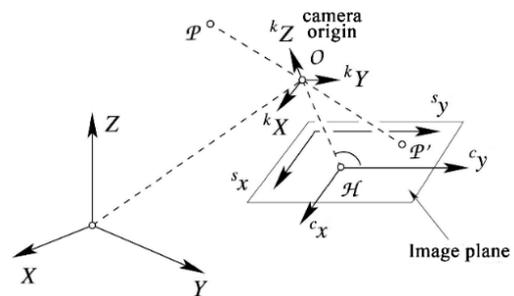


**Figure 1.** Flow chart of the procedure illustrated in this paper for crack metric evaluation.

### 2.1 Working reference systems

A crucial step is to define the working reference systems. As widely known in photogrammetry and computer vision, an object in the Object Reference System (ORS, identified with capital letter  $XYZ$  in Figure 2) is mapped with a rigid-body transformation into the Camera Reference System (CRS,  ${}^kX{}^kY{}^kZ$ ), then to the image plane ( ${}^cX{}^cY$ ) by a central projection and finally to the sensor ( ${}^sX{}^sY$ ) by an affine transformation (Förstner et al., 2016).

Concerning the ORS choice, two different approaches are considered. When facing the problem of crack opening in RC structure, it may be enough to estimate crack geometry, and thus to scale the photogrammetric model, by measuring one (or many) distance in real world and on images. Therefore, when no



**Figure 2.** Scheme of the coordinate reference systems involved in the projection process. Letters ( $XYZ$ ) denote the object reference system; ( ${}^kX{}^kY{}^kZ$ ) is the camera reference system; ( ${}^cX{}^cY$ ) is the image coordinate system and ( ${}^sX{}^sY$ ) is the sensor reference system (image adapted from Förstner et al., 2016).

Ground Control Points (GCPs) are provided, the ORS is fixed to the CRS of the first camera ( ${}^kX{}^kY{}^kZ$  in Figure 2), with the  $Z$  axis pointing towards the camera viewing direction. This reference system will be hereafter called Local Camera Object Reference System (LCORS). On the other hand, if a periodical monitoring of the crack opening is required, referencing the photogrammetric model in a permanently materialized World Reference System (WRS) becomes mandatory. This can be either a local cartesian system or a cartographic reference system. To this end, a similarity transformation from the LCORS to the WRS must be estimated based on a set of GCPs.

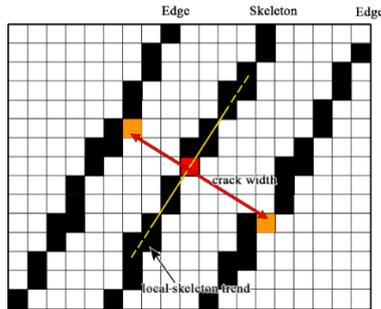
To define a general and flexible workflow, this paper proposes to work in a LCORS for stereo-pair orientation and crack 3D reconstruction. Afterwards, the model can be either scaled or georeferenced in an WRS, depending on the aims of the specific work. For the latter case, a total station can be employed to derive GCPs world coordinates on hardly accessible bridges. Alternatively, a UAV-based photogrammetric model of the whole structure, with a sufficiently high image Ground Sample Distance (GSD), can be built with a traditional SfM workflow (Pinto et al, 2020). Distinctive features can be extracted from the photogrammetric model and used as GCPs.

### 2.2 Crack Detection

Before crack detection, images are converted to grayscale and undistorted, to remove non-linear distortions and work with an affine camera model (Förstner et al., 2016). To improve binarization results, images are pre-processed by using histogram equalization and bi-dimensional Wiener filter (Kumar et al., 2010). Moreover, a mask is manually drawn on the images to roughly identify the cracked area and help the binarization process. The masked images are then binarized with the adaptive Bradley's method (Bradley et al., 2007), which uses a locally adaptive thresholds for each pixel by computing the local mean intensity around the pixel neighbourhood. Morphological operators, such as closing and filling operators, are then used to reduce noisy crack pixels and identifying a realistic crack shape, by exploiting pixel connectivity proprieties (Tanaka et al., 1998).

To derive crack skeleton, a skeletonization algorithm based on the medial axis transform (Lee, 1982) was employed. Finally, the crack edges were extracted by using the Canny algorithm (Canny, 1986) on binarized images. Since cracks edges are detected on binarized images, the different edge detection algorithms, such as Canny, Sobel, or Prewitt (Maini et al., 2009) provide comparable results.

Once skeleton and edges are detected, for each skeleton pixel, the closest edge pixels are searched, as it will be necessary for computing the width of the crack (Figure 3). To this end, for each skeleton pixel, a local linear trend is computed on the skeleton chain within a moving  $5 \times 5$  research window. The two correspondent edge pixels are searched along the direction perpendicular to the local skeleton trend.



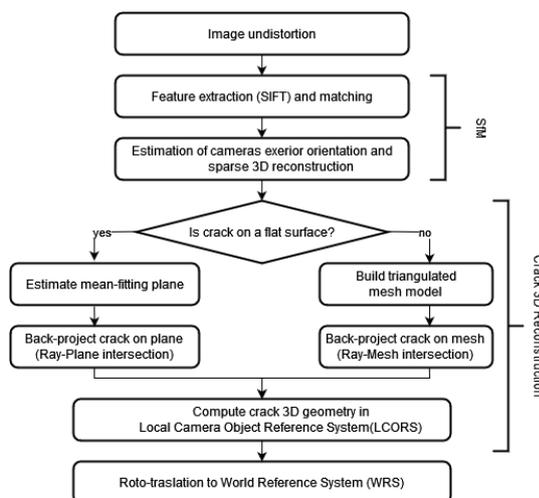
**Figure 3.** Scheme of the modelled crack geometry, with the central skeleton chain and the edges. For each point of the skeleton, the two corresponding edge pixels are searched on the perpendicular direction to the local skeleton trend.

### 2.3 SfM and 3D scene reconstruction

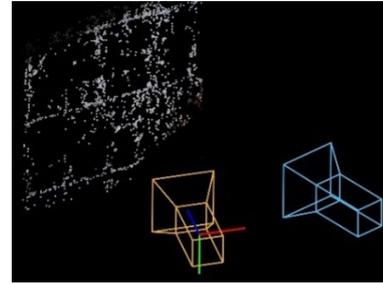
To estimate cameras relative orientation and reconstruct a 3D model of the scene, a traditional SfM approach was pursued. The main steps of the procedure are illustrated in Figure 4. Interesting points and their descriptors were extracted by Scale-Invariant Feature Transform (SIFT) operator (Lowe, 2004), thanks to its well known robustness to scale, rotations, perspective and illumination changes. To this end, the Matlab toolbox VLFEAT (Vedaldi et al., 2010) was employed. Matches between interesting points were then selected by minimizing the squared Euclidean distance between descriptors. False matches were rejected by exploiting the coplanarity constrain given by:

$$\mathbf{x}'^T \mathbf{E} \mathbf{x}'' = 0 \quad (1)$$

where  $\mathbf{E}$  is the Essential Matrix (as the camera interior orientation is known),  $\mathbf{x}'$  and  $\mathbf{x}''$  are the vectors of the normalized image coordinates of the homologous points, respectively on the first and on the second image. The  $\mathbf{E}$  matrix was estimated from homologous points by a five-points method (Stewénius et al., 2006), combined with RANSAC (Fischler and Bolles 1981) to improve robustness.



**Figure 4.** Flow chart of SfM and crack reconstruction process.



**Figure 5.** Camera exterior orientation and sparse point cloud. The axis XYZ of the LCORS are marked respectively with red, green and blue segment.

Based on the estimated  $\mathbf{E}$  matrix, the location and orientation of the second camera in LCORS was solved iteratively by least-squares, by using the Computer Vision Toolkit developed by Fusiello (Fusiello, 2018) (Figure 5). The location of the second camera was determined up to a scale-factor, which is estimated by using one (or many) distance in real world. The 3D coordinates in LCORS of homologous points were reconstructed by triangulation, producing a sparse point cloud.

### 2.4 Crack back-projection on the 3D model

Given a continuous 3D model of the object (e.g., wall, bridge beam or pier) and crack image coordinates on one image, crack 3D coordinates can be derived by intersecting the projective rays with the model. If a crack lays on a flat surface, this can be modelled as plane, and crack 3D coordinates can be determined through a ray-plane intersection (Liu et al., 2016). For non-flat surfaces (e.g., bridge piers), a triangulated mesh model must be used, and a ray-mesh intersection problem must be solved (Liu et al., 2019). The ray-plane intersection method is easier and of faster computation, but it clearly works only on flat surfaces. The ray-mesh intersection method has a wider applicability, but it requires the construction of a triangulated mesh model of the object. In this work, both methods were tested.

#### 2.4.1 Crack back-projection by ray-plane intersection

Given a straight-line preserving perspective camera, the projection of the world point  $\mathbf{X}$  in homogeneous coordinates to the point  $\mathbf{x}$  on the sensor is given by (2) (Förstner et al., 2016):

$$\mathbf{x} = \mathbf{K}[\mathbf{R} | \mathbf{t}] \mathbf{X} = \mathbf{P} \mathbf{X} \quad (2)$$

where  $\mathbf{R}$  is the camera rotation matrix and  $\mathbf{t}$  is the camera translation vector.  $\mathbf{K}$  is the calibration matrix, containing the interior orientation parameters

$$\mathbf{K} = \begin{bmatrix} c & cs & u_0 \\ 0 & c(1+m) & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

where  $c$  is the principal distance,  $u_0$  and  $v_0$  are the coordinates of the principal point,  $m$  and  $s$  are respectively scale difference and shear parameters. The projection matrix  $\mathbf{P}$  can be written as

$$\mathbf{P} = \mathbf{K}[\mathbf{R} | \mathbf{t}] = [\mathbf{P}_{1:3} | \mathbf{P}_4] \quad (4)$$

As the projective mapping is not invertible, with one camera only and starting from (2), it is possible to derive

$$\mathbf{X} = \mathbf{X}_0 + \lambda(\mathbf{K}\mathbf{R})^{-1} \mathbf{x} \quad (5)$$

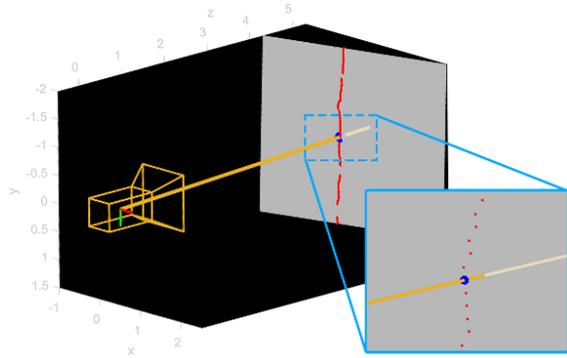


Figure 6. Schematic view of the ray-plane intersection

that describe the direction of the projective ray from the camera perspective centre  $\mathbf{X}_0$  to the 3D point  $\mathbf{X}$ . To compute the world coordinates of  $\mathbf{X}$ , more images are required, and the problem is solved by triangulation. Alternatively, the 3D point  $\mathbf{X}$  can be determined from just one image as the intersection of the perspective ray, starting from  $\mathbf{X}_0$  and passing through the projection  $\mathbf{x}$  on the image, with the actual surface of the object. In most bridge beams or abutments, the surface on which cracks lay can be modelled as a plane. Thus, the problem may be reduced in finding  $\mathbf{X}$  as the intersection of the projective ray with the wall surface plane (Figure 6). The projective ray can be written as a line, in Plücker's coordinates, passing through the camera perspective centre and the infinite-far away point  $\mathbf{X}_\infty$  along the direction  $\lambda(\mathbf{KR})^{-1}\mathbf{x}$ . The camera projection centre is given by the null space of the projection matrix  $\mathbf{P}$  as:

$$\mathbf{X}_0 = -\mathbf{P}_{1:3}^{-1}\mathbf{P}_4 = -(\mathbf{KR})^{-1}\mathbf{Kt} \quad (6)$$

It should be recalled that given two arbitrary points  $\mathbf{V}$  and  $\mathbf{W}$  with coordinates:

$$\mathbf{V} = [V_1 V_2 V_3 V_4]^T = [\mathbf{V}_0 V_h]^T \quad (7)$$

$$\mathbf{W} = [W_1 W_2 W_3 W_4]^T = [\mathbf{W}_0 W_h]^T \quad (8)$$

(where  $V_h$  and  $W_h$  denote the homogenous parts of the vectors, while  $V_0$  and  $W_0$  are the inhomogenous or Euclidean parts), the line  $\mathbf{L}$  passing through them can be written in Plücker's coordinates as:

$$\mathbf{L} = [L_1 L_2 L_3 L_4 L_5 L_6]^T = \begin{bmatrix} \mathbf{L}_h \\ \mathbf{L}_0 \end{bmatrix} = \begin{bmatrix} V_h W_0 - W_h V_0 \\ \mathbf{V}_0 \times \mathbf{W}_0 \end{bmatrix} \quad (9)$$

The two points are the projective centre  $\mathbf{X}_0$  and the infinite-far away point  $\mathbf{X}_\infty$ . Moreover, considering a generic plane  $\mathbf{A}$  with homogenous coordinates:

$$\mathbf{A} = [A B C D]^T = [\mathbf{A}_h A_0]^T \quad (10)$$

the coordinates of the intersection point  $\mathbf{X}$  are then given by:

$$\mathbf{X} = \mathbf{L} \cap \mathbf{A} = \begin{bmatrix} A_0 \mathbf{I} & -S(\mathbf{A}_h) \\ -\mathbf{A}_h^T & \mathbf{0}^T \end{bmatrix} \quad (11)$$

where  $S(\mathbf{A}_h)$  is the skew-symmetric matrix induced by the vector  $\mathbf{A}_h$  (Förstner et al., 2016). Computing the intersection of all the crack points (edges and skeleton) detected on one image with the mean-fitting plane of the wall surface, a scaled and geometric-consistent 3D point cloud of the crack is obtained.

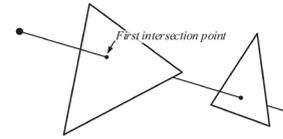


Figure 7. Ray-triangle intersection (credits: Liu et al., 2019)

#### 2.4.2 Crack back-projection ray-triangle intersection

If the cracked surface is not flat, this must be modelled with a triangulated mesh model, and a ray-mesh intersection problem must be solved (Figure 7). To this end, the algorithm proposed by Möller et al. (1997), and implemented in the *TriangleRayIntersection* Matlab toolkit developed by Jarek Tuszynski, was employed. Requirements are a triangulated mesh surface, the ray origin and direction. The camera centre of projection  $\mathbf{X}_0$  is the ray origin; the direction of the projective ray is described by  $\lambda(\mathbf{KR})^{-1}\mathbf{x}$ . If more than one intersections are found, only the first intersection point is taken (Figure 7).

Despite the ray-mesh intersection approach was tested to derive a general workflow, in this paper the focus will be mostly on the ray-plane intersection method. Many bridge beams or abutments, in fact, can be modelled as a plane, and the method is simpler and of faster computation.

#### 2.5 Crack width metric estimation

Once the coordinates of the crack edges and skeleton are derived in the LCORS, the crack width is estimated. To this end, for each point of the skeleton, the 3D Euclidean distance between the corresponding two edge points was computed (Figure 8). The correspondence of skeleton and edges points was established on the images, as described in Section 2.2, and it clearly holds also in the LCORS.

The described procedure needs just one image to compute the crack width, as the 3D coordinates of the crack points are derived by intersecting the projective rays with a plane or a mesh. Nevertheless, a stereo-pair of images is required to compute the external orientation of the cameras (see Section 2.4) and the point cloud from which the plane or the mesh is built. Therefore, the crack edge and skeleton points can be back-projected from each of the two cameras to overcome occlusion problems, and to check the crack reconstruction consistency. Overall, the image taken with the image plane more parallel to the wall surface is considered as the more reliable.

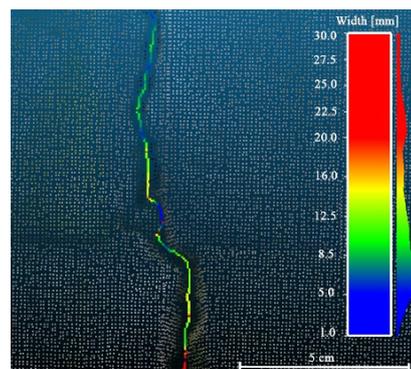
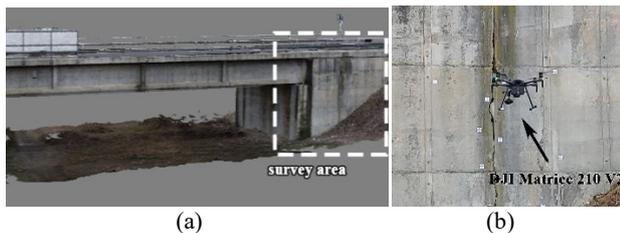


Figure 8. 3D crack skeleton reconstructed by the ray-plane intersection method. For each skeleton point, the colorscale represents the crack width computed as the 3D Euclidean distance between the closest edge points.

### 3. THE CASE STUDY

#### 3.1 Experiment setup and instruments

As test field for the proposed method, the Ponte di Tuna bridge on river Trebbia was selected (Figure 9a). The bridge was located between the Provincial Roads n. 28 and n. 40 south of Piacenza (44°57'51" N, 9°35'26" E - Italy). It was 684 m long and crosses the Trebbia River. It consisted of 19 spans of 36 m. The bridge deck was set on 5 PC beams with a height of 2.10 m, which support a 0.2 m-height RC slab. The bridge was built in 1980 and it was renovated in 2011.



**Figure 9.** (a) Ponte Tuna used for the validation test. (b) The crack present on the West abutment.

A wide crack present on the West bridge abutment (Figure 9b) was chosen as case study for this paper. The choice was motivated by the easy accessibility of the area, which allowed several markers to be placed on the wall surface for error assessment. A local cartesian was materialized by a multi-station Leica MS60, and 16 plastic targets, attached with a double-side tape on the wall, were measured with millimetric accuracy (Fagandini et al., 2017).

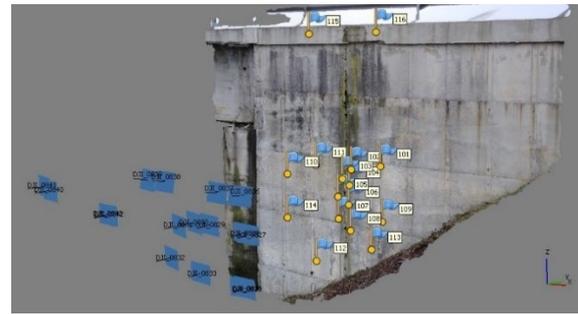
To acquire crack images, a commercial quadcopter DJI Matrice 210 V2 was used. It was equipped with a DJI Zenmuse X5S camera with 20Mpx 4/3" CMOS sensor and a DJI MFT 15mm/1.7 ASPH lens. The camera was mounted on a 3-axis gimbal. Several images of the crack were acquired by manually flying the UAV at a distance of  $\sim 4$  m from the abutment surface, to have a Ground Sample Distance (GSD) of  $\sim 1$  mm. Two slightly convergent images, image A taken from the left-hand side and image B taken from the right-hand side (Figure 10), were then selected for testing the procedure.



**Figure 10.** Stereo-pair of images used in the test: (a) image A (left-hand side); (b) image B (right-hand side). The two images are not undistorted yet.

#### 3.2 Camera calibration and image distortion correction

Before processing the images for detecting the cracks and reconstructing the scene, non-linear lens distortions must be removed from the images. To this end, the camera was pre-calibrated by setting up a calibration-field with 12 GCPs, placed on ground and measured with millimetric accuracy by a multi-station Leica MS60. A set of 48 convergent images was acquired by flying over the calibration field (Ioli et al., 2021).



**Figure 11.** Reference photogrammetric model of the bridge abutment built by Agisoft Metashape. Blue rectangles are the oriented cameras, while the flags denote the GCPs.

The procedure was similar to the well-known checkerboard calibration method (Zhang, 2000), but it allowed for maintaining an image scale comparable to that kept during the actual bridge survey.

#### 3.3 Reference photogrammetric model

A reference photogrammetric model was built by using the commercial SfM software Agisoft Metashape 1.8.1. To this end, 17 images, acquired from different point of view (Figure 11), were oriented by solving a Bundle Block Adjustment (BBA) based on 12 Ground Control Points (GCPs), with a prior accuracy of 1 mm. The same camera interior orientation parameters, estimated as in Section 3.2, were used as initial values in the BBA. Due to the high number of GCPs available and a rather good acquisition geometry, a self-calibration was carried out (James et al., 2020).

The quality of the photogrammetric block was assessed based on the 4 remaining targets, used as Check Points (CP), resulting in a global RMSE = 0.7 mm, evenly distributed in the three directions. A photogrammetric dense point cloud was then built by using full-resolution images (i.e., *Ultra-high Quality* parameter in Metashape), and *mild* depth maps filtering, resulting in a point spacing comparable to the average GSD of the images (i.e.,  $\sim 1$  mm). This photogrammetric dense cloud will be the reference for assessing the estimated crack width.

#### 3.4 Crack edge and skeleton detection

As first step of the workflow, images were undistorted, by using the calibrated interior orientation, to correct non-linear distortions. Afterwards, the algorithm described in Section 2.2 was employed to detect crack edges and skeleton on both image A and image B. Moreover, for each skeleton pixel on both the images, the corresponding edge pixels were searched along the perpendicular direction to the skeleton. The crack skeleton chain and edges detected on image A are marked in Figure 12. The algorithm for detecting crack skeleton and edges well performed for most of the crack, both where the crack is significantly large ( $width > 1$  cm) and for thinner cracks ( $3$  mm  $<$  width  $<$  1 cm). As the average GSD of the images was  $\sim 1$  mm, the thinnest detectable crack width was considered as  $3 \times GSD = 3$  mm (Liu et al., 2019). However, in some areas such as those represented in the magnified cut-out windows 4 and 5 of Figure 12, the crack skeleton and edges were not properly recognized. In the area of cut-out 4, the algorithm failed because of the presence of some small plants grown into the crack. In the area of cut-out 5, on the other hand, the crack had some very thin areas ( $\leq 1$  mm), where the edge-contrast was weak, and thus the binarization algorithm broke the skeleton and edge chains.



**Figure 12.** Crack skeleton pixel chain (red dots) detected from image A, superimposed to the undistorted RGB image A. In the five magnified cut-outs views (1-5), the detected edges and skeleton are represented respectively with green and red dots.

The results obtained from image A and image B were overall comparable and the crack skeleton and edges were successfully detected in most of the images. On image A, 80% of the crack length was detected, while this percentage slightly reduced to 73% on image B. Some discrepancies in crack edge detection between image A and image B occurred, mostly limited to few pixels, and located in tough areas, such as those with vegetation growth in the crack, with very small crack breath, or with shadows or occlusions. These differences may clearly lead to differences in crack reconstruction, when back-projecting crack information from the two images to the 3D model. To overcome this problem, the image acquired with an optical axis more perpendicular to the wall surface (i.e., image B, in this case study) will be considered as the most reliable, as occlusions or shadows issues are less severe.

### 3.5 3D crack reconstruction

Image A and image B were then oriented relatively by procedure illustrated in Section 2.3, and the sparse point cloud was computed by triangulating homologous points in the LCORS. The model was scaled by using the distance between two targets, measured by the multi-station Leica MS60, with millimetric accuracy. To speed up the procedure, no dense point cloud was computed at this stage, as it was considered not strictly necessary if enough and homogeneously distributed homologous points were detected on the image frames.

A global mean-fitting plane was fitted on the entire point cloud and the image coordinates of the crack edge and skeleton were back-projected on this plane, by applying the method discussed in section 2.4.1. However, as a RC wall surface has usually some roughness and it is not perfectly homogeneous, a global mean-fitting plane may be a poor surface approximation for the whole crack. Therefore, a second iteration was carried out: for each point of the back-projected crack skeleton, a subset of the point cloud was extracted by looking for its 30 nearest neighbours, based on the Euclidean distance. A local-mean-fitting plane was fitted again on this point cloud subset, and skeleton and edges image points were back-projected again on this local-mean-fitting plane. The whole procedure was repeated for both image A and image B. Finally, once crack and skeleton edges points were back-projected in the LCORS, the crack width was estimated for each point of the skeleton, by computing the 3D Euclidean distance between the two correspondent edge points.

		From image A			From image B		
		X	Y	Z	X	Y	Z
R-P	Mean [mm]	0.3	0.4	-0.1	0.1	0.4	-0.1
	Std [mm]	1.1	1.3	0.8	0.8	1.3	0.6
	RMSE [mm]	1.2	1.4	0.8	0.8	1.4	0.6
R-M	Mean [mm]	0.2	0.1	-0.2	0.2	0.1	-0.1
	Std [mm]	0.8	0.5	0.8	0.9	0.5	0.6
	RMSE [mm]	0.9	0.5	0.9	0.9	0.5	0.7

**Table 1.** Statistics of the differences between the back-projected coordinates of 10 targets with those measured on the field by the multistation Leica MS60. The back-projection method is marked with acronym R-P for ray-plane intersection, and R-M for ray-mesh intersection.

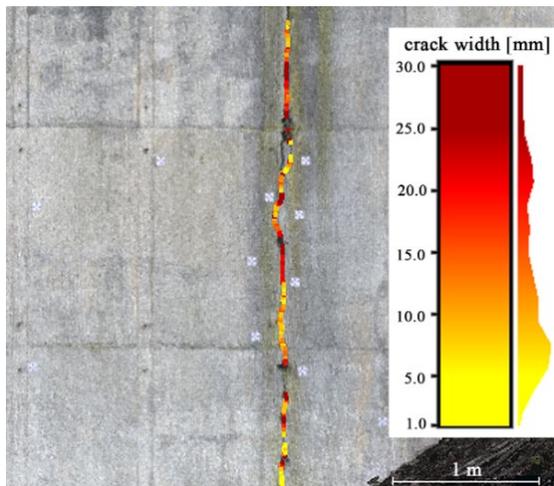
To assess the ray-plane intersection procedure, targets placed on the wall surface were employed. A Helmert transformation from the LCORS to the WRS was estimated based on 4 targets (global RMSE computed on the remaining 10 targets of 0.6 mm). The ray-plane intersection procedure was carried out in the WRS, starting from the image coordinates of the remaining 10 targets. The estimated 3D coordinates were then compared with those measured by the multi-station Leica MS60 and the results are listed in Table 1. The RMSE was in the order of the millimetre for the 3 coordinates, which was comparable with the images GSD and with the accuracy of the reference measurements obtained with the multi-station. Moreover, no significant differences between the back-projected coordinates of the targets obtained from image A and B were found.

Additionally, the ray-mesh intersection method (see Section 2.4.2) was tested. To build the triangular mesh model, the software Agisoft Metashape was employed. Image A and image B were imported in Agisoft Metashape, maintaining the cameras exterior and interior orientation fixed to those obtained by roto-translating the cameras from the LCORS to the WRS (i.e., the same configuration as that used to validate the ray-plane intersection method was kept). The mesh model was derived directly by computing depth maps on full resolution images, and then exported to Matlab again. The *TriangleRayIntersection* algorithm, developed by Jarek Tuszynski, was employed to derive the 3D coordinates of the 10 targets, starting from image coordinates of image A and image B. The differences between the estimated coordinates and those measured by the multi-station Leica MS60 are listed in Table 1. As it could be expected, the result of the ray-mesh intersection procedure is slightly better than that obtained by the ray-plane intersection method (sub-millimetric RMSE for ray-mesh method), but still, it is in line with both the image GSD and accuracy of the target measurements. It can be concluded that if the wall surface is a plane, the two methods give comparable results. The ray-mesh intersection method, however, becomes mandatory when cracks lay on non-planar surfaces.

### 3.6 Crack width estimation and validation of the results

The crack width was estimated as described in Section 2.5, and associated to each skeleton point, so that it can be easily visualized as a 3D point cloud (Figure 13).

To validate the results of the crack width estimation, the 3D point clouds of skeleton and edges were roto-translated into WRS by a Helmert transformation estimated based on 4 targets (see Section 3.5). The point clouds were then compared with the reference photogrammetric dense cloud obtained by Agisoft Metashape (see Section 3.3), by using the open-source software Cloud Compare (Cloud Compare, 2022).



**Figure 13.** 3D skeleton point, colored based on the estimated crack width. The skeleton points are superimposed to the photogrammetric point cloud as reference. Next to the colorbar, the histogram shows the distribution of the estimated width values.

To quantify the width estimation error, 11 skeleton points were selected, and the crack width was manually measured on the photogrammetric reference cloud. Among the 11 validation points, 5 were selected where the crack width was ranging between 3 mm and 1 cm (i.e., *narrow crack*), while the remaining 6 where the crack width was between 1 cm and 2 cm (i.e., *wide crack*). The statistics of the differences, computed for narrow and wide cracks separately, are summarized in Table 2. The global RMSE was around 1 mm, that is in line with the average image GSD.

	Image A		Image B	
	<i>Narrow</i>	<i>Wide</i>	<i>Narrow</i>	<i>Wide</i>
Mean [mm]	0.7	1.0	0.7	0.7
Std [mm]	0.6	0.7	0.4	0.9
RMSE [mm]	0.9	1.2	0.8	1.1

**Table 2.** Statistics of the differences between the estimated crack width, respectively from image A and image B, with manual measurements obtained from the reference photogrammetric point cloud on the same positions.

Measurements are divided in two groups: *Narrow* group refers to narrow crack segments with  $3\text{ mm} < \text{width} < 1\text{ cm}$ ; *Wide* refers to wide crack segments with  $1\text{ cm} < \text{width} < 2\text{ cm}$ .

#### 4. CONCLUSIONS

This paper presents an image-based procedure for metric reconstruction of cracks in RC/PC structures, by acquiring stereo-pair of images with UAVs, enabling the inspection of large infrastructures such as highway bridges or viaducts.

Cracks are identified on images by using image binarization and edge detection techniques, by exploiting pixel crack connectivity to identify the skeleton. Crack edges are further detected with edge-detection algorithms. This well-established method allows for a good identification of crack features in most of the image (between 70% and 80% of the crack length, in the performed study case), with crack misdetection in areas with shadows or occlusions, vegetation growth within the crack or with very thin breadth.

UAV-based stereo-pair of images, from different point of view are employed to build a photogrammetric model by SfM. The

model is built up to a similarity transformation. To retrieve metric information from the model, at least one (accurate) distance measurement is required, but no GCPs are in principle needed. To georeference the RC damage on a structure or to set up an image-based monitoring system to observe the evolution in time of a crack, GCPs are required. These can be acquired either with a total station or from an existing UAV photogrammetric model of the whole structure, which makes the procedure suitable for large infrastructure monitoring.

Crack 3D metric reconstruction is then carried out by back-projecting crack skeleton and edges, detected on one image, on a 3D model. This can be either a plane, when dealing with planar wall surfaces, or a triangular mesh model. When convergent images are used, the one with the image plane more parallel to the wall surface should be considered as the most reliable for skeleton and edges detection and crack 3D reconstruction.

The procedure was tested in a field investigation on a large crack present on a bridge abutment. The choice was driven by the possibility to place on the structure several reference target for conducting an error analysis. Images were acquired with a camera mounted on a quadcopter, with an average GSD of 1 mm (flying at a distance of  $\sim 4\text{ m}$  from the abutment surface). The results of the test highlighted that cracks with breadth up to three times the average GSD were properly reconstructed with millimetric accuracy, which was comparable to the average GSD of the images. If smaller cracks have to be detected with sub-millimetric accuracy, longer focal lengths and shorter working distances are required.

Future development of the procedure consists of improving crack detection methods, e.g., by exploiting the use of properly trained CNN, such as ResNet, for identifying cracks on images on a larger image dataset, with limited operator supervision. Moreover, more than two images may be used for sparse and dense reconstruction through SfM. This would allow for acquiring both convergent images, for improved scene reconstruction robustness, and images parallel to the wall surface for better crack skeleton and edges detection. Finally, the procedure may be extended, including additional surface structural defects (e.g., steel corrosion, spalling, humidity stains) to be detected on images.

#### REFERENCES

- Abdel-Qader, I., Abudayyeh, O., Kelly, M.E., 2003. Analysis of Edge-Detection Techniques for Crack Identification in Bridges. *J. Comput. Civ. Eng.*, 17(4), 255–263.
- ARTBA 2022. Bridge report. <https://artbabridgereport.org> (28 March 2022).
- Bradley, D., Roth, G., 2007. Adaptive Thresholding using the Integral Image. *J. Graph. Tools*, 12(2), 13–21.
- Canny, J., 1986. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-8(6), 679–698.
- Cha, Y.J., Choi, W., Büyüköztürk, O., 2017. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Comput. Civ. Infrastruct. Eng.*, 32(5), 361–378.

- Chaize, P., Dagbert, M., 2019. Sécurité des ponts: éviter un drame. Rapport d'information. <https://www.senat.fr/rap/r18-609/r18-6091.pdf> (28 March 2022).
- Cloud Compare, 2022. 3D point cloud and mesh processing software Open-Source Project <https://cloudcompare.org/> (28 March 2022).
- Fagandini, R., Federici, B., Ferrando, I., Gagliolo, S., Pagliari, D., Passoni, D., Pinto, L., Rossi, L., Sguerso, D., 2017. Evaluation of the Laser Response of Leica Nova MultiStation MS60 for 3D Modelling and Structural Monitoring. *Lecture Notes in Computer Science.*, 10407, 93–104.
- Fischler, M.A., Bolles, R.C., 1981. Random sample consensus. *Commun. ACM*, 24(6), 381–395.
- Förstner, W., Wrobel, B., 2016: *Photogrammetric Computer Vision*. Springer Nature, Cham
- Fusiello, A., 2018. *Visione computazionale. Tecniche di ricostruzione tridimensionale*. Franco Angeli, Milano.
- Hackl, J., Adey, B.T., Woźniak, M., Schümperlin, O., 2018. Use of Unmanned Aerial Vehicle Photogrammetry to Obtain Topographical Information to Improve Bridge Risk Assessment. *J. Infrastruct. Syst.*, 24(1).
- Hutchinson, T.C., Asce, M., Chen, Z., 2006. Improved Image Analysis for Evaluating Concrete Damage. *J. Comput. Civ. Eng.*, 20(3), 210–216.
- Jahanshahi, M.R., Kelly, J.S., Masri, S.F., Sukhatme, G.S., 2009. A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. *Struct. Infrastruct. Eng.*, 5(6), 455–486.
- James, M.R., Antoniazza, G., Robson, S., Lane, S.N., 2020. Mitigating systematic error in topographic models for geomorphic change detection: accuracy, precision and considerations beyond off-nadir imagery. *Earth Surf. Process. Landforms*, 45(1), 2251–2271.
- Kim, B., Cho, S., 2019. Image-based concrete crack assessment using mask and region-based convolutional neural network. *Struct. Control Heal. Monit.*, 26(8).
- Kim, H., Lee, J., Ahn, E., Cho, S., Shin, M., Sim, S.H., 2017. Concrete Crack Identification Using a UAV Incorporating Hybrid Image Processing. *Sensors*, 17(9), 2052.
- Kim, H., Sim, S.H., Spencer, B.F., 2022. Automated concrete crack evaluation using stereo vision with two different focal lengths. *Autom. Constr.*, 135, 104136.
- Kumar, S., Kumar, P., Gupta, M., Nagawat, A.K., 2010. Performance Comparison of Median and Wiener Filter in Image De-noising. *Int. J. Comput. Appl.*, 12(4), 27–31.
- Ioli, F., Pinto, L., Ferrario, F., 2021. Low-cost DGPS assisted aerial triangulation for sub-decimeter accuracy with non-RTK UAVs. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B2-2021, 25–32.
- Lee, D.T., 1982. Medial Axis Transformation of a Planar Shape. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-4(4), 363–369.
- Liu, Y.F., Cho, S., Spencer, B.F., Fan, J.-S., 2016. Concrete Crack Assessment Using Digital Image Processing and 3D Scene Reconstruction. *J. Comput. Civ. Eng.*, 30(1), 04014124.
- Liu, Y.F., Nie, X., Fan, J.S., Liu, X.G., 2019. Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction. *Comput. Civ. Infrastruct. Eng.*, 35(5), 511–529.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, 60(2), 91–110.
- Maini, R., Aggarwal, H., 2009. Study and comparison of various image edge detection techniques. *Int. J. Image Process.*, 3(1), 11
- Möller, T., Trumbore, B., 1997. Fast, Minimum Storage Ray-Triangle Intersection. *J. Graph. Tools*, 2(1), 21–28.
- Pinto, L., Bianchini, F., Nova, V., Passoni, D., 2020. Low-cost UAS Photogrammetry for Road Infrastructures' Inspection. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B2-2020, 1145–1150.
- Rao, A.S., Nguyen, T., Palaniswami, M., Ngo, T., 2021. Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure. *Struct. Heal. Monit.*, 20(4), 2124–2142.
- Rau, J.Y., Hsiao, K.W., Jhan, J.P., Wang, S.H., Fang, W.C., Wang, J.L., 2017. Bridge crack detection using multi-rotary UAV and object-base image analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W6, 311–318.
- Stewénius, H., Engels, C., Nistér, D., 2006. Recent developments on direct relative orientation. *ISPRS J. Photogramm. Remote Sens.*, 60(4), 284–294.
- Tanaka N, Uematsu K., 1998. A Crack Detection Method in Road Surface Images Using Morphology. MVA 1998 IAPR Workshop on Machine Vision Application, 17-19.
- Tuszynski J., 2022. Triangle/Ray Intersection, Version 1.7, MATLAB Central File Exchange. <https://www.mathworks.com/matlabcentral/fileexchange/33073-triangle-ray-intersection> (28 March 2022).
- Vedaldi, A., Fulkerson, B., 2010. VLFeat-An open and portable library of computer vision algorithms. *Proc. Int. Conf. Multimed.*, 1469-1472. doi.org/10.1145/1873951.1874249
- Zhang, Z., 2000. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11), 1330–1334.