Automated Inspection within Galleries of Large Dams

Helena LAASCH*, Nathalie RYTER*, Isabelle STEFFEN*, Alexander RESKE, Ephraim FRIEDLI, Valens FRANGEZ, Switzerland
*These authors contributed equally to this work

Key words: Dam Inspection, Crack Detection, Indoor Mobile Mapping, DSLM Camera, Laser Scanning, Depth Camera

SUMMARY
Dams are widely used for energy production and need to be inspected to ensure their stability and most importantly security. So far, the inspection within dam galleries is done periodically in a manual manner, thus making it time-consuming, labor-intensive, and subjective. Alternatively, automated inspections can be challenging due to the narrow galleries, sparse lighting, and texture of the concrete. This paper presents a case study on data acquisition and evaluation of three different sensors concerning data suitability for automated crack detection within such environments. The evaluated sensors in this study are *Leica* mapping system *BLK2GO*, the *Lumix* digital single-lens mirrorless (DSLM) camera *DMC-FZ2000*, and the *time-of-flight depth camera Helios Lucid*. The measurements and data acquisition were done using a mobile robotic platform in a 220 m long dam in the Swiss Alps.

This paper proposes a data evaluation pipeline for the extraction of georeferenced cracks. The processing outcomes are critically analyzed in terms of geometric and prediction accuracies, the repeatability of the predicted cracks, as well as the economic aspects of proposed measurement solutions. For the crack detection, different methods based on convolutional neural networks (CNN) were assessed using the datasets acquired with the three sensors. The DSLM images showed the best outcome due to their highest resolution, allowing the detection of cracks wider than 1 mm. Furthermore, an approach for locating cracks within the 3D digital model of the dam was proposed. For this georeferencing task, two approaches are presented, namely the *BLK2GO* trajectory and the photogrammetric model obtained from the DSLM images. Ultimately, the findings indicate that using a combination of sensors outperforms stand-alone solutions, i.e., a combination of the DSLM for the acquisition of images georeferenced based on the *BLK2GO* trajectory. This paper is a contribution toward inspection task automation within industrial environments using digital sensing and robotic technologies.

Automated Inspection within Galleries of Large Dams

Helena LAASCH*, Nathalie RYTER*, Isabelle STEFFEN*, Alexander RESKE, Ephraim FRIEDLI, Valens FRANGEZ, Switzerland
*These authors contributed equally to this work

1 INTRODUCTION

To ensure the stability of dams, the inspection of cracks and their shape changes over time have to be carried out periodically every few years. So far, this activity is still done manually by skilled professionals, who walk through the galleries and annotate cracks and their changes over time. This is time-consuming, labor-intensive, performed in challenging work conditions, and involves the subjectivity of a skilled professional. Therefore, finding an automated sensor-based alternative is of high interest. Until now, a few approaches for automated inspection employing mobile platforms and sensors were published in the literature. E.g., Gehring et al. [1] proposed to use a mobile platform carrying different sensors to perform inspection of an offshore platform. Phung et al. [2] conducted bridge inspection using a 3D model of the bridge for the pathfinding of a camera-equipped unmanned aerial vehicle (UAV). Khaloo et al. [3] proposed using a photogrammetric model created from images taken by a manually controlled UAV equipped with a camera to detect flaws and defects on the outside of a gravity dam. The work of Sarker et al. [4] focused in particular on the suitability of depth cameras for crack detection. This published work indicates that there is a strong interest in automating such inspections, however according to author's knowledge the potential is not yet fully exploited and further developments are needed, as will be presented in this contribution.

We made use of the following three sensors, namely the *Leica* mapping system *BLK2GO* [5], the *Lumix* system camera *DMC-FZ2000* [6], and the *Lucid* depth camera *Helios Time of Flight 3D camera* [7] for data collection in a dam environment. The sensors were mounted on a wheeled robot to allow for autonomous inspection. This paper proposes a workflow taking in as an input the data collected with sensors, i.e. images and point clouds, which results in 3D coordinates of detected cracks. Within this study, two different convolutional neural networks (CNN) are tested to automate the crack detection.

The focus of this paper is on assessing the suitability of the three sensors for the inspection of galleries of large dams. The results obtained with the sensors are compared with each other concerning the geometrical and prediction accuracy of cracks, repeatability of the results, and economic aspects. The desired outcome for each sensor is to determine the smallest detectable crack with an absolute positioning accuracy at a centimeter-level. Their performance was tested in an arch-gravity dam in the Swiss Alps. The environment is challenging due to the sparse lighting conditions, the homogeneous texture of the wall, the narrow gallery of the long dam, and the absence of the global navigation satellite system (GNSS) signal.

The data collection, the measurement environment, and the used instruments are presented in Chapter 2. Chapter 3 presents the methods used for extracting and locating the cracks from the collected data. In Chapter 4, the results of the different sensors are presented and compared with each other. Finally, the paper concludes with a summary of the main findings and an outlook in Chapter 5.

2 FIELDWORK AND PREPROCESSING

2.1 Measurement Environment

The measurement site was an arch-gravity dam in the Swiss Alps. The gallery of the dam was around 220 m long and its diameter was about 1.80 m. Nine consoles installed in the gallery served as ground control points (GCPs) and are part of a permanently installed geodetic network to monitor the dam and build the dam coordinate system. The overall geometry of the network is not optimal for measurements since the shape of the tunnel is long and narrow, meaning the GCPs are almost on a line. The environment is challenging due to the sparse and non-homogeneous lighting condition, the variable humidity, and the homogeneous, repetitive pattern on the surface of the concrete (see Figure 1).

Figure 1: Conditions inside the gallery of the arch-gravity dam

2.2 Experimental Setup

2.2.1 Sensors Used in the Assessment

In the context of this paper, three sensors (see Figure 2) were selected based on their functionalities and were tested in terms of their suitability for detecting cracks, i.e. an all-in-one solution, a sensor acquiring high resulted images, and an active sensor providing 3D point information.

Figure 2: Sensors to be tested (from left to right): Leica BLK2GO [8], Lumix DMC-FZ2000 [9], Lucid Helios Time of Flight 3D camera [10]

The *Leica BLK2GO* is used as the all-in-one solution, high-resolution images were acquired with the *Lumix DMC-FZ2000*, and the depth images were captured with the *Lucid Helios Time of Flight 3D camera*. We refer to the mentioned sensors in the following parts as BLK2GO, DSLM, and depth camera, respectively. To facilitate the comparison of the three sensors, selected specifications have been collected in Table 1.

Table 1: Selected sensor specifications

	BLK2GO	DSLM	Depth Camera
Output	Point cloud, images, Trajectory information	RGB images or videos	Depth and intensity images
Sensors	LiDAR, IMU, 4 Cameras	1-inch MOS	0.5-inch Sony DepthSense IMX556 CMOS
Localization	GrandSLAM	n/a	n/a
Relevant Acquisition Frequency	LiDAR: 420'000 pts/s	Images: n/a Videos: 25 fps	15 Hz
Pixel Size at 1 m Distance	Image: 1.6 mm	0.26 mm	1.7 mm
Operation Range	LiDAR: 0.5 – 25 m	n/a	0.3 m – 6.0 m
Image Resolution	4.8 MP	20.1 MP	0.3 MP
Field of View	300° x 135°	4° x 3° - 73° x 53° (focal length 8.8 – 176 mm)	59º x 45º

The *Leica BLK2GO* is a handheld imaging laser scanner. This mapping solution is based on the GrandSLAM technology which makes use of an IMU, visual and light detection and ranging (LiDAR) simultaneous localization and mapping (SLAM) and is used for the positioning of the *BLK2GO*. Besides a point cloud and a trajectory, the system also provides for each image a timestamp, position, and orientation with respect to the mapping frame. Since it is a one-button solution, it is not possible to change any acquisition settings of the cameras or the LiDAR. The digital single-lens mirrorless (DSLM) camera *DMC-FZ2000* from *Lumix* is a system camera that can acquire images and videos. The latter can then be used to extract the individual image frames. The depth camera *Helios Time-of-Flight* by *Lucid Vision Labs* is an active sensor that provides depth and intensity images. The depth camera is controlled by a Python script on an external computer, where the recorded data is stored. An external power source is also required for the operation. The most important parameters adjusted in our case were exposure time, operating distance, averaging over frames, and gain.

2.2.2 Mobile Platform

A four-wheeled robot, the *Super Mega Bot* from *Inspectorbots* and customized by the *RobotX Center for Robotics*, was used to carry the sensor. The robot was used to assess the applicability of such a vehicle in the dam environment. The sensors mentioned in Chapter 2.2.1 were mounted on the platform, as shown in Figure 3. Thus, their relative positions with respect to each other stayed the same during the experiments.

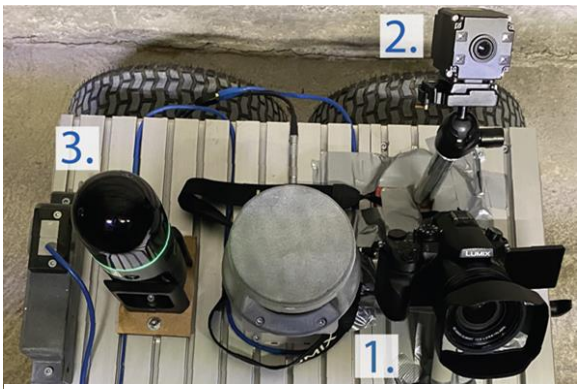

Figure 3: Sensors (1. DSLM, 2. depth camera, 3. BLK2GO) mounted on the Super Mega Bot

2.3 Calibration

User-calibrations of the sensors were carried out before acquiring the data. For the *BLK2GO*, we followed the guided user-calibration as instructed by the manufacturer *Leica Geosystems* [11]. To correct for the lens distortions in the images acquired with the DSLM, we used a black and white chessboard. The acquisition of the board was done at different angles and distances and the images were used for the estimation of the radial distortion and intrinsic parameters of the camera using the *Camera Calibrator* app of the toolbox *Image Processing and Computer Vision* from *MATLAB* [12]. Using these parameters we undistorted the images in the following processing steps. The intensity images acquired with the depth camera are impacted by the scanning distance and the angle of incidence as a consequence of the measurement configuration, as well as the camera related inter-pixel variations. In order to remove these effects, the intensity images were used to estimate the corresponding reflectance images by applying a previously estimated calibration function. For the latter, an independent experiment was set up using a flat highly-diffuse white wall to acquire depth and intensity images following the procedure described in Frangez et al. [13].

2.4 Data Collection

2.4.1 Off-Site Experiments

Before carrying out the measurements on-site in the arch-gravity dam, we conducted experiments to determine the optimal acquisition settings for all the sensors used. We assessed the DSLM to determine whether videos or images are more suitable as capture mode. Our experiments showed that videos are less dependent on the speed of the robot since the number of frames extracted from the videos is flexible and can be selected by the operator, and thus in post-processing the desired overlap of the images can be chosen. The following on-site experiments were therefore executed in the video mode.

To achieve optimal lighting of the images, the appropriate camera settings had to be determined experimentally. This means the optimal balance between the focal length, the exposure time, and the ISO value had to be determined. The darker the environment, the higher the exposure time or a higher ISO value should be set. If the exposure time is longer, the image gets blurrier due to movement and with a higher ISO value, the image exhibits higher noise. Finally, the focal length had to be set, affecting how much light passes through the lens. However, a short focal length lets more light pass through, which leads to a lower focal depth. Experimentally chosen parameters for our particular setup and environment were ISO 800, focal length 3.5 mm, and exposure time 1/60 s.

A similar process of settings selection was chosen for the depth camera. However, since this is an active sensor, the acquisition is independent of the ambient lighting conditions. Due to the dynamic measurements, the averaging over the acquired frames was not used and the exposure time was set to the lowest one possible, i.e. 250 µs.

Since no settings can be changed on the *BLK2GO*, the measurement conditions, i.e. the surrounding lighting had to be adjusted accordingly to obtain data of sufficient quality. We therefore carried out tests with different light conditions using various lighting setups. The results of our investigation indicated that bright and homogeneous light distribution within the field-of-view of the camera provides the optimal results. Obtaining images of sufficient quality

also has a direct impact on the trajectory estimation, since the visual SLAM directly relies on the images and the features that can be identified from the images. Furthermore, the bright and homogeneous lighting conditions also improve the quality of images captured by DSLM. Certain kind of AC lighting (i.e. fluorescent or incandescent light sources) can cause a pulsating effect within depth camera measurements, therefore we avoided using such light sources within this study. Finally, the optimal solution for all three sensors was using a construction site lamp (LED with 50 W power and 4000 lm).

2.4.2 Reference Data

To assess the data acquired in the arch-gravity dam with the three sensors, we used the *Leica* laser scanner *RTC360* [14] and the total station *Leica TS60* [15] to collect the reference data. Both datasets were expressed in the dam geodetic network by using the GCPs, which were either equipped with targets for scanning or prisms for tachymetry measurements. The *RTC360* was used to scan the entire gallery and the registration was done using black and white checkerboard targets distributed in the gallery with approximately 1 m spacing. The *TS60* was used to measure reflectorless the locations of distinctive points along the cracks in a part of the dam. The positions of the cracks could be determined on a millimeter-level accuracy.

2.4.3 On-Site Experiments

So-called velocity, and repeatability tests as well as 360°-coverage acquisition of the predefined gallery sections were carried out on-site to determine the suitability of the three sensors for crack detection. The velocity test was conducted by acquiring the data with the mobile robot moving at velocities of approximately 0.25 m/s, 0.5 m/s, and 1 m/s. We found that higher-quality images can be acquired by moving more slowly, i.e. approximately 0.25 m/s. Within the repeatability tests we aimed at finding whether the same crack can be detected within different measurement epochs.
Since the sensors were fixed on the robot, their field of view direction was kept constant. To cover both sides of the gallery, the robot had to move back and forth along the gallery, switching the orientation at the end. If the coverage of the ceiling is desired, the sensors on the platform had to be re-oriented. Therefore, to ease the acquisition we performed 360°-coverage acquisition of the gallery sections manually using the DSLM and depth cameras.

3 METHODS AND PROCESSING

3.1 Reference Data

The registration and georeferencing of the scans acquired with the *RTC360* were carried out in *Leica Cyclone* [16]. The registered point cloud with the achieved geometrical accuracy of 2 mm was considered sufficient as reference data for the application presented herein. The coordinates of the cracks measured with the total station *TS60* were exported directly as georeferenced coordinates from the instrument and were used in the subsequent analysis.

3.2 Crack Detection Algorithms

Two different CNNs were used to extract cracks from the acquired datasets, namely for the intensity images we used a CNN model from *Amberg Infra 7D* [17], and for the RGB images we used a self-trained U-Net-based architecture. According to [17], the *Amberg Infra 7D* model

was trained with intensity images obtained by profile laser scanning data of tunnels. For the CNN model training we used a combination of images containing cracks acquired in different environments (i.e. provided by *AXPO* [18]) as well as images acquired on-site in the dam gallery. Combinations of the optimizers adam and stochastic gradient descent (SGD) [19], and the loss functions lovasz, focal, dice, and binary cross-entropy (BCE) [20] were used for the model training. The model with the adam optimizer and the focal loss function resulted in only few false-positive values but detected only wider cracks. The model with the adam optimizer and the BCE loss function had many false-positive values but also detected thin cracks. The model with the adam optimizer and the dice loss function was a compromise between the previous two models. To predict the cracks, a combination of these three best-performing models was used. The cracks were predicted with each model and the resulting binary images were overlaid and merged. The result was a grayscale image with four shades of gray used for the number of models that predicted a crack at a pixel-level (see Figure 4 left). This combined image represents the probability of the existence of a crack. For the final predicted crack image, a majority voting was performed. Less probable pixels, which are only predicted with one model are labeled as a crack when they belong to an area which already contains some pixels that are detected by two or more models. The resulting image (see Figure 4, right) reduced the number of false-positive detected cracks compared to the image predicted with only one model.

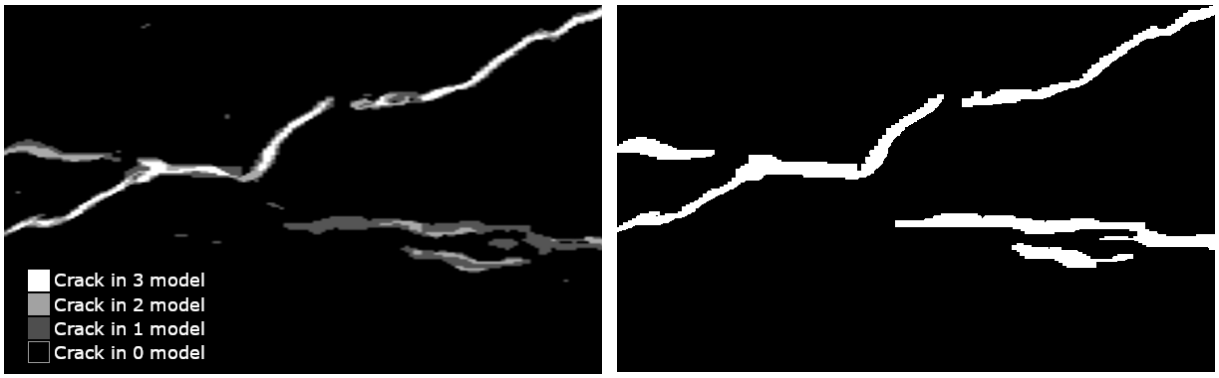

Figure 4: Cumulative image of the three selected crack detection models (left) and the resulting image from the combination of the three models (right)

3.3 Trajectory

To georeference each detected crack, the image where the crack was identified needs to be georeferenced by knowing the camera extrinsic parameters. The latter is known if a trajectory of a sensor and subsequently of the robot is known. This is the case for the *BLK2GO* system. The images of the *BLK2GO* are in the mapping frame like the corresponding acquired point cloud. By knowing the transformation parameters of the point cloud, the trajectory, and the extrinsic parameters of the *BLK2GO* camera, the images were transformed in the dam coordinate system.

Due to the challenging environment for the SLAM positioning system, the trajectory experienced a non-linear drift. To reduce the impact of the drift, the point cloud was cut into 15 – 20 m sections. A separate rigid transformation was performed by computing the transformation between two targets within a particular *BLK2GO* point cloud section and their corresponding targets within the reference data of the *RTC360* point cloud expressed in the dam

coordinate system. The transformation was then applied to the whole respective point cloud section. Due to the underdetermination of the transformation, the roll along the axis between two targets is not well constrained. The roll is however expected to be small because of the internal IMU.

Figure 5 and Equations 1 and 2 show how the transformation between the *BLK2GO* and the dam coordinate system was done. First, the *BLK2GO* coordinate system was translated and rotated into the first target point so that the x-axis pointed in the direction of the second point. After scaling in the x-direction to reduce the drift, the rotation and translation into the dam coordinate system were performed.

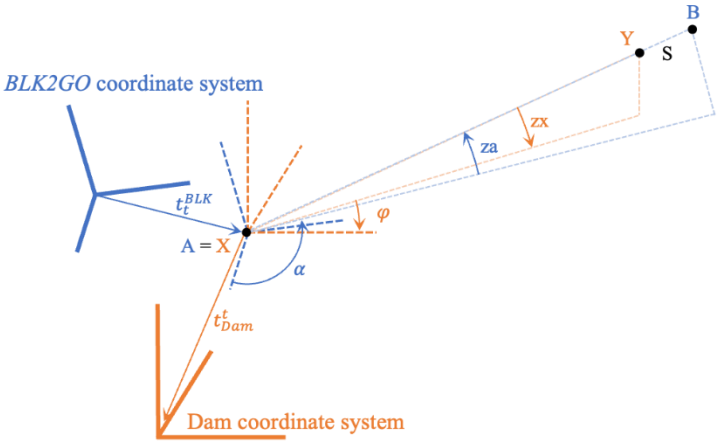

Figure 5: Transformation between BLK2GO and dam coordinate system

In the next step, we estimate the trajectory of the DSLM and the depth camera based on the known constant offset between the sensors and *BLK2GO*. For simplicity, the following is shown only for the DSLM, however it can be also applied to the depth camera setup. In order to calculate the offset between the DSLM and the *BLK2GO*, an image of the same scene was acquired from each sensor, with the robot being stationary. These images were referenced with the use of the *RTC360* point cloud. Distinctive points were manually measured in the images and the point cloud. Out of these corresponding points, the extrinsic camera parameters were calculated with the direct linear transformation (DLT) [21]. Figure 6 gives an overview of all coordinate systems and their known transformations.

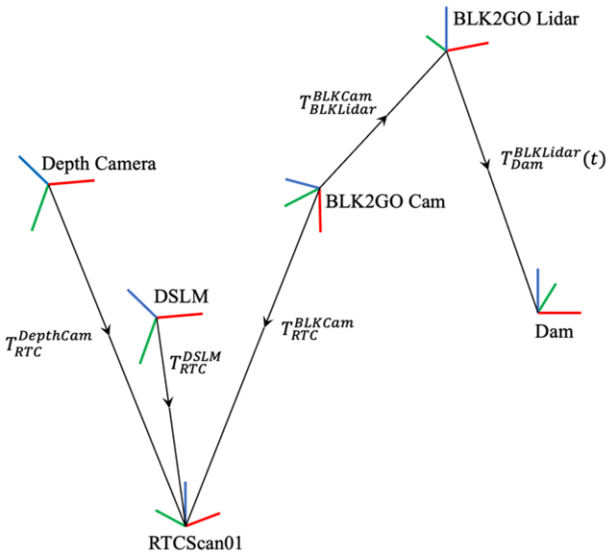

Figure 6: Different coordinate systems and their known transformation matrices used in the data

$$R_{Dam}^{BLK} = R_z(\varphi) \cdot R_x(zx) \cdot S_x \cdot R_x(za) \cdot R_z(\alpha) \quad (1)$$

$$t_{Dam}^{BLK} = R_{Dam}^{BLK} \cdot t_t^{BLK} + t_{Dam}^t \quad (2)$$

where:

$$t_t^{BLK} = -A,\ t_{Dam}^t = X,\ \varphi = arctan\ \frac{Y_y - X_y}{Y_x - X_x},\ zx = arctan\ \frac{Y_z - X_z}{\|X-Y\|},\ s = \frac{\|X-Y\|}{\|A-B\|},$$

$$za = rctan\ \frac{B_z - A_z}{\|A-B\|},\ \alpha = arctan\ \frac{B_y - A_y}{B_x - A_x}$$

The transformations from both sensor systems to the *RTC360* coordinate system T_{RTC}^{DSLM}, T_{RTC}^{BLKCam} are known through the DLT. The transformation $T_{BLKLidar}^{BLKCam}$ from the *BLK2GO* camera to the *BLK2GO* trajectory is given from the calibration of the instrument. $T_{Dam}^{BLKLidar}(t)$, which describes the transformation from the *BLK2GO* trajectory coordinate system to the dam coordinate system, was given by georeferencing the trajectory for each timestamp. Finally, Equation 3 shows the entire transformation from the DSLM coordinate system to the dam coordinate system.

$$T_{Dam}^{DSLM} = T_{Dam}^{BLKLidar}(t) \cdot T_{BLKLidar}^{BLKCam} \cdot (T_{RTC}^{BLKCam})^{-1} \cdot T_{RTC}^{DSLM} \quad (3)$$

To find the corresponding *BLK2GO* trajectory point for a DSLM image, time synchronization between the sensors is required. For this purpose, the experimentally determined beginning of the movement was determined from the different sensor images and the time offsets were calculated.

3.4 Georeferencing Cracks with Trajectory

With the trajectory and the georeferenced point cloud we were able to estimate the 3D coordinates for each pixel of an image containing the detected cracks. From the intrinsic parameters of the camera and the known orientation, the vector between the camera center and the pixel detected as crack could be calculated which is the line of sight (LOS). With a nearest neighbor search perpendicular to the LOS approximate 3D coordinates for the pixel detected as a crack could be determined at the centimeter-level. To reduce computation time only the border pixels of the crack were used to calculate the 3D coordinates.

3.5 Georeferencing Cracks Using a Photogrammetric Model

Another approach to georeference the images is using a photogrammetric model was explored as well. The image overlap of about 80% was large enough to estimate a photogrammetric model using *Agisoft Metashape* [22]. The model was georeferenced by choosing at least three georeferenced points in the images. It was possible to export the camera positions and orientations to perform the crack detection (see Chapter 3.3) and the crack projection onto a point cloud (see Chapter 3.4).

3.6 Segmentation

In order to know which projected pixels belong to which crack, the 3D coordinates were clustered with DBSCAN [23]. The parameters were defined experimentally, i.e. the radius was set to 0.05 m, and 5 points were required to be within the search radius. To assign pixel regions that correspond to the same crack acquired in different epochs we calculated several parameters to allow for correct crack matching. For this particular task we had chosen a bounding box oriented according to the axes of the dam coordinate system, its diagonal and center point, as well as the number of points in order to describe the crack as unique.

4 Results and Discussion

4.1 Crack Detection

We were able to successfully detect cracks in the images of all sensors (see Table 3 and Figure 9). Figure 7 shows the number and width of detected cracks for each sensor type. In the DSLM

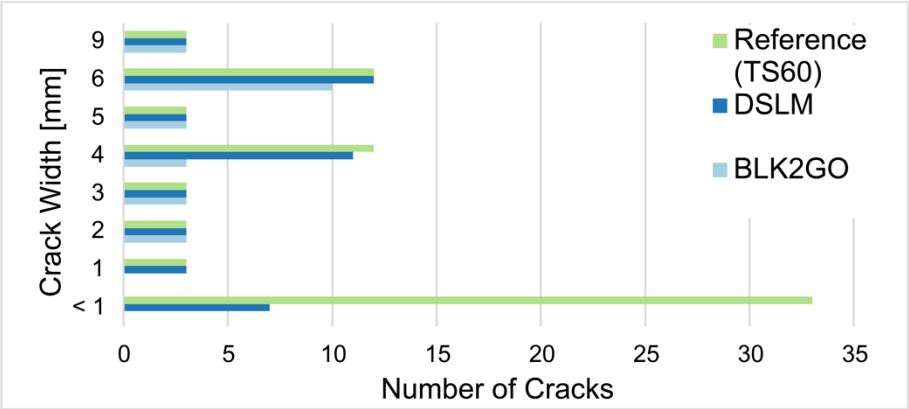

Figure 8: Number of detected cracks (in one image per section per run) of both sensors compared with the number of reference cracks in dependency of their widths

Table 2: Comparison between a subsampled image, tiled image without subsampling, and an orthophoto of the DSLM sensor with detected

	Section 1		
	%	– –	– +
DSLM subsampled	80	4	13
DSLM tiled	72	4	133
DSLM orthophoto	69	3	95

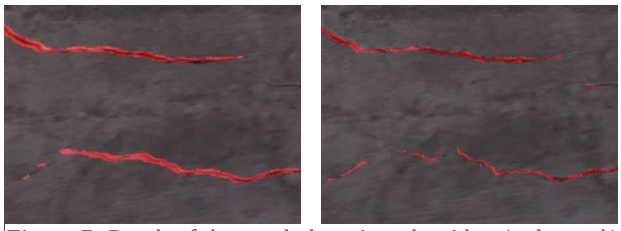

Figure 7: Result of the crack detection algorithm (red: crack) with a subsampled image (left) and with a tiled image (right)

images, 97 % of the cracks wider than 1 mm could be detected by the self-trained CNN. Most of the cracks wider than 2 mm were detected in the *BLK2GO* images. To avoid running out of memory, we tiled the images, i.e. we cut the images used in the prediction step into tiles of size 896 x 896 pixels (see Table 2 and Figure 8). The disadvantage of this tilling is that cracks are more interrupted, wide cracks are not always detected, and the number of false-positive values is higher. In the generated orthophotos from the photogrammetric model, cracks are detected but not to their full extent.

Table 2 shows the percentage (%) of the crack lengths that were detected within the detection algorithm. Besides from the percentage values, very important indicators for the performance of the CNN are the numbers of false-negative predictions (– –) and false-positive predictions (– +). The false-negative predictions should be avoided, and false-positive predictions should be minimized, meaning, it is better to detect a crack that is not actually a crack, rather than not detecting a crack at all. Comparing the cracks detected using the orthophoto with respect to the other two DSLM results (i.e. using subsampled and tiled datasets), the percentage of detected cracks is lower. Additionally, the false-negative values are lower by 10%. For the intensity images acquired with the depth camera, the crack detection algorithm was able to predict cracks that are wider than 6 mm, however about 60% of the crack was not detected (see Figure 9).

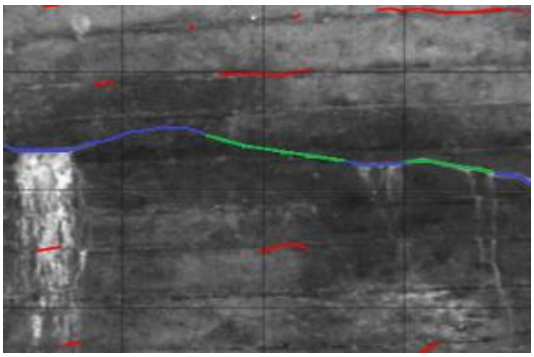

Figure 9: Cracks on the intensity image of the depth camera (green: correct detected, blue: false-negative values, red: false-positive values)

To use the proposed method for crack detection, repeatability of the crack prediction must be ensured. This part of the analysis was performed using three measurement epochs (runs) acquired in three different parts of the gallery (see Chapter 2.4.3). Ideally, the detected cracks should have the same size and form in all runs used in the analysis. The evaluation results based on a single image are shown in Table 3. The higher the value and more similar the values are between the epochs per section, the higher the repeatability. The datasets acquired with the *BLK2GO* have more false-negative values than the DSLM, and the percentage of the detected crack lengths is higher using the DSLM. This can be traced back to the images of lower-resolution of the *BLK2GO*, as well as the sensor settings which are not adjustable to be able to acquire images with higher quality. Generally, there is higher agreement between epochs acquired with the DSLM than those acquired with the *BLK2GO*.

Table 3: Detected crack lengths (%), false-negative predictions (– –), and false-positive predictions (– +) for one image in each section and run

	Section 1						**Section 2**						**Section 3**					
	DSLM			BLK2GO			DSLM			BLK2GO			DSLM			BLK2GO		
Run	%	– –	– +	%	– –	– +	%	– –	– +	%	– –	– +	%	– –	– +	%	– –	– +
1	80	4	13	37	8	32	90	1	17	54	6	26	66	3	20	80	6	26
2	79	4	23	31	8	25	91	2	18	47	6	28	75	3	19	74	6	25
3	86	5	12	39	8	32	81	3	6	58	6	21	77	2	11	58	5	15

4.2 Accuracy of the Geometry

The accuracy of the trajectory and the point cloud directly influence the accuracy of the 3D coordinates of detected cracks. As described in Chapter 3.3, the *BLK2GO* trajectory indicates a large drift due to the challenging gallery shape, i.e. elongated tunnel. The impact of the drift on the point cloud estimation is shown in comparison to the *RTC360* point cloud (see Figure 10). The two point clouds were aligned based on the calculation of the transformation between the GPCs in the *RTC360* dataset and the *BLK2GO* dataset (see Chapter 3.3).

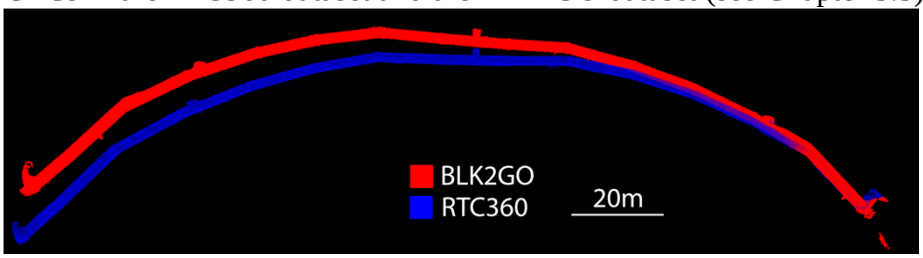

Figure 10: Indication of a drift of the BLK2GO point cloud compared to the RTC360 point cloud, which served as ground truth (top view)

The impact of the drift of the trajectory onto the detected cracks is shown in Figure 11. The figure shows extracted cracks of nine consecutive DSLM images, which appear shifted with respect to each other, on a level of a few centimeters. Furthermore, the distances between the georeferenced *BLK2GO* targets and the targets of the *RTC360* point are shown in Table 4. The deviations increase from the

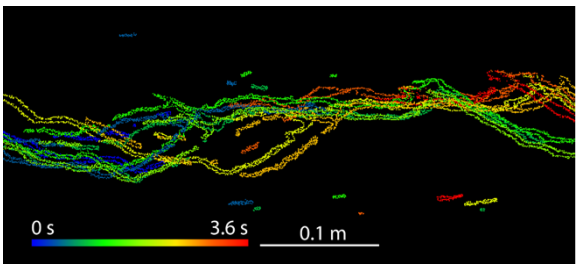

Figure 11: Projected cracks from nine consecutive images (Δt=0.4 s)

start of the gallery, i.e. part used for alignment, to the end, where they reach several centimeters. To avoid influences of our special setup on the trajectory, the *BLK2GO* was carried by hand in run four. The deviations of run four are in a similar range as the others which leads to the assumption that the setup has no impact on the quality of the trajectory. The cloud-to-cloud (C2C) distances between the *RTC360* and *BLK2GO* point clouds are also shown in Table 4. The C2C differences are smaller since the profile of the gallery does not change much. As the drift is along the axis of the gallery, not the whole impact of the drift on the point cloud can be shown with the C2C distances.

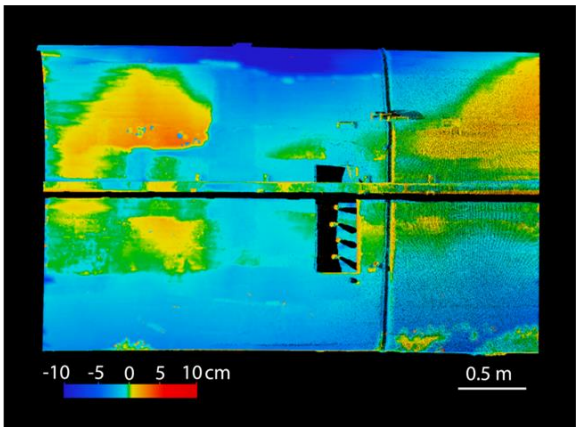

Figure 12: C2M comparison between photogrammetric model and RTC360 point

In the case where the georeferencing is done with a photogrammetric model (see Figure 12) the accuracy of the cloud-to-mesh (C2M) distance is higher with a mean of 2.8 cm and a standard deviation of 2.4 cm, as compared to the C2C distance of the *BLK2GO* point cloud (see Table 4). The achieved accuracy is in the centimeter range, therefore it is considered sufficiently good for this application. One of the disadvantages of using the photogrammetric model is that the computation takes much more time than using the trajectory. Also, if the *BLK2GO* point cloud is partitioned into smaller sections, which are then individually transformed, the drift of the trajectory could be significantly reduced.

Table 4: Mean (μ) and standard deviation (σ) of the target distance and cloud-to-cloud (C2C) distance between RTC360 and BLK2GO point cloud (all values in [cm])

	Section 1				**Section 2**				**Section 3**			
	Target distance		C2C		Target distance		C2C		Target distance		C2C	
Run	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
1	3.2	2.0	0.7	0.7	27.2	8.4	0.9	2.8	23.0	13.6	1.8	5.7
2	4.0	2.0	0.6	0.8	17.2	9.0	1.6	2.8	36.8	26.5	2.3	6.1
3	9.5	7.5	0.8	0.8	34.7	13.5	1.5	3.9	31.1	13.1	1.5	5.3
4	3.4	1.9	2.4	5.3	58.2	19.9	4.2	5.5	21.6	11.8	2.3	2.1

5 CONCLUSION

We carried out a suitability investigation of three sensors, namely *Leica* mapping system *BLK2GO*, the *Lumix* digital single-lens mirrorless (DSLM) camera *DMC-FZ2000*, and the *time-of-flight* depth camera *Helios Lucid*, for the task of autonomous inspection of cracks in dam galleries. The data was collected in an arch-gravity dam in the Swiss Alps. The crack extraction from the acquired data was done based on convolutional neural networks (CNN). A summary of the advantages and disadvantages of each sensor for crack detection application is outlined in Table 5.

Overall, the crack extraction results obtained with the DSLM camera are the most satisfactory, given suitable ambient conditions, i.e. sufficient and homogeneous lighting. If DSLM is used as a standalone solution, computation of a photogrammetric model needs to be carried out for localization, which can be computationally intensive. However, if data processing, as well as crack extraction, can be done in post-processing and not in real-time, this disadvantage becomes irrelevant. We showed that the geometry of the photogrammetric model is accurate on a centimeter-level, which is considered enough for the application presented herein and the orthophoto derived therefrom yields promising results concerning the number of false-negative cracks. For the given measurement environment and the selected processing steps, the *BLK2GO* data cannot be considered suitable due to being prone to drifts as a consequence of the homogeneous elongated geometry of the dam. Also, the depth camera performance was not satisfactory, since the amount of successfully extracted cracks from the intensity images was the lowest compared to the other two sensors. The limited performance using intensity images was likely related to their relatively low resolution.
To further improve the optimal solution using the DSLM camera, future work should focus on improving the CNN-based crack detection algorithms, which should ideally be trained with more images containing cracks of various shapes and sizes. Also, to increase the versatility, images acquired in various lighting conditions and surface types (e.g. wet or dry concrete surfaces, which are visually different) should be included. This might as well improve the repeatability of the crack extraction, which would increase the reliability of the automatic inspection. Furthermore, practical aspects of choosing a more suitable platform, e.g. a drone, for data acquisition should be considered to improve the flexibility (i.e. to ensure adjustable field-of-view of the sensor to achieve full coverage of the gallery walls) and to ensure a constant distance to the gallery wall. The selected suitable platform should be equipped with a sufficient lighting source to ensure homogeneous light coverage on the wall.

Table 5: Summary of the advantages and the disadvantages of the evaluated sensors

	BLK2GO	DSLM	Depth Camera
Pros	- cracks larger than 2 mm can be detected - images, trajectory information, point cloud in one measurement - images aligned with the point cloud - not post-processing intensive	- cracks larger than 1 mm can be detected - low number of false-negative cracks - low-cost sensor	- cracks larger than 6 mm* can be detected - active sensor, i.e. no additional lighting needed - low-cost sensor * not evaluated with smaller cracks
Cons	- passive sensor, i.e. needs additional lighting - large drift of trajectory - which varies throughout the point cloud - high number of false-negative cracks - no acquisition setting adjustment possible	- photogrammetric model or trajectory with point cloud needed for the crack localization - passive sensor, i.e. needs additional lighting - localization is computationally expensive	- photogrammetric model or trajectory with point cloud needed for the crack localization - external battery and storage - needed - localization is computational expensive

6 ACKNOWLEDGMENTS

The authors would like to thank T. Zimmer and P. Riesen (*AXPO Holding AG*) for making this project possible and for providing valuable support. We thank P. Spohn and M. Burri (*Amberg INFRA 7D*) for the help with the crack detection algorithm using intensity images.

REFERENCES

[1] C. Gehring, P. Fankhauser, L. Isler, R. Diethelm, S. Bachmann, M. Potz, L. Gerstenberg and M. Hutter, "ANYmal in the Field : Solving Industrial Inspection of an Offshore HVDC Platform with a Quadrupedal Robot," in *Field and Service Robotics*, Singapore, 2021.

[2] M. D. Phung, T. H. Dinh, V. T. Hoang and Q. Ha, "Automatic Crack Detection in Built Infrastructure Using Unmanned Aerial Vehicles," in *Proceedings of the 34th International Symposium on Automation and Robotics in Construction (ISARC)*, Taipei, Taiwan, 2017.

[3] A. Khaloo, D. Lattanzi, A. Jachimowicz and C. Devaney, "Utilizing UAV and 3D Computer Vision for Visual Inspection of a Large Gravity Dam," *Frontiers in Built Environment*, vol. 4, 2018.

[4] M. M. Sarker, T. A. Ali, A. Abdelfatah, S. Yehia and A. Elaksher, "A COST-EFFECTIVE METHOD FOR CRACK DETECTION AND MEASUREMENT ON CONCRETE SURFACE," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vols. XLII-2/W8, 2017.

[5] Leica Geosystems, "Leica BLK2GO," 2022. [Online]. Available: https://shop.leica-geosystems.com/blk2go. [Accessed 07 01 2022].

[6] Panasonic Deutschland, "DMC-FZ2000 Superzoom Kamera," 2022. [Online]. Available: https://www.panasonic.com/de/consumer/foto-video/dmc-fz2000.html. [Accessed 07 01 2022].

[7] LUCID Vision Labs Inc., "Helios Time of Flight (ToF) 3D Camera," 2021. [Online]. Available: https://thinklucid.com/helios-time-of-flight-imx556/. [Accessed 07 01 2022].

[8] Leica Geosystems AG, [Online]. Available: https://shop.leica-geosystems.com/sites/default/files/inline-images/BLK2GO_Side_Logo_Light_WBackground.png. [Accessed 29 04 2022].

[9] Foto Wolf. [Online]. Available: https://www.foto-wolf-dresden.de/images/product_images/original_images/055_FY2016_FZ2000_vorne.jpg. [Accessed 29 04 2022].

[10] fabrimex Systems, [Online]. Available: https://www.fabrimex-systems.ch/686-large_default/lucid-helios-tof-time-of-flight-camera.jpg. [Accessed 29 04 2022].

[11] Leica Geosystems AG, "Leica Geosystems," 2022. [Online]. Available: https://leica-geosystems.com/. [Accessed 07 01 2022].

[12] The MathWorks Inc., "Camera Calibrator," 2022. [Online]. Available: https://ch.mathworks.com/help/vision/ref/cameracalibrator-app.html. [Accessed 07 01 2022].

[13] V. Frangez, D. Salido-Monzú and A. Wieser, "Surface finish classification using depth camera data. In Automation in Construction (Vol. 129, p. 103799).," *Elsevier BV*, 2021.

[14] Leica Geosystems AG, "Leica RTC360 3D Laser Scanner," 2022. [Online]. Available: https://leica-geosystems.com/products/laser-scanners/scanners/leica-rtc360. [Accessed 07 01 2022].

[15] Leica Geosystems AG, "Leica Nova TS60," 2022. [Online]. Available: https://leica-geosystems.com/products/total-stations/robotic-total-stations/leica-nova-ts60. [Accessed 07 01 2022].

[16] Leica Geosystems AG, "Leica Cyclone 3D Point Cloud Processing Software," 2022. [Online]. Available: https://leica-geosystems.com/products/laser-scanners/software/leica-cyclone. [Accessed 07 01 2022].

[17] Amberg Infra 7D, "Amberg Infra 7D," 2022. [Online]. Available: https://amberggroup.com/infra-7d/. [Accessed 07 01 2022].

[18] Axpo Holding AG, "Axpo," [Online]. Available: www.axpo.com. [Accessed 07 01 2022].

[19] Torch Contributors, "TORCH.OPTIM," 2019. [Online]. Available: https://pytorch.org/docs/stable/optim.html. [Accessed 07 01 2022].

[20] Torch Contributors, "TORCH.NN.FUNCTIONAL," 2019. [Online]. Available: https://pytorch.org/docs/stable/nn.functional.html. [Accessed 07 01 2022].

[21] E. Dubrofsky, "Homography Estimation," The University of British Columbia, Vancouver, 2009.

[22] Agisoft, "Discover intelligent photogrammetry with Metashape," 2022. [Online]. Available: https://www.agisoft.com/. [Accessed 07 01 2022].

[23] scikit-learn developers, "sklearn.cluster.DBSCAN," 2022. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html. [Accessed 07 01 2022].

BIOGRAPHICAL NOTES

Helena Laasch, Nathalie Ryter and Isabelle Steffen are MSc students in Geomatic Engineering at *ETH Zürich*. Alexander Reske is a research engineer at the Robotic Systems Lab, *ETH Zürich*. Ephraim Friedli is an engineering specialist in geodesy at *AXPO Holding AG*, Switzerland. Valens Frangez is a scientific assistant and a doctoral student in the group of Geosensors and Engineering Geodesy, *ETH Zürich*.

CONTACTS

Helena Laasch, Nathalie Ryter, Isabelle Steffen
Institute of Geodesy and Photogrammetry
ETH Zürich
SWITZERLAND
Email: laaschh@ethz.ch, rytern@ethz.ch, isteffen@ethz.ch