Depth-camera-based rebar detection and digital reconstruction for robotic concrete spraying

Valens Frangez¹ · Ena Lloret-Fritschi² · Nizar Taha² · Fabio Gramazio² · Matthias Kohler² · Andreas Wieser¹

Received: 8 April 2021 / Accepted: 5 October 2021 / Published online: 16 October 2021
© The Author(s) 2021

Abstract

In this paper, we tackle the challenge of detection and accurate digital reconstruction of steel rebar meshes using a set of industrial depth cameras. A construction example under investigation in this paper is robotic concrete spraying, where material is sprayed onto double-curved single layered rebar meshes. Before the spraying process can start, the location and geometry of the rebar mesh needs to be accurately know. We present an automatic image-based processing approach of depth images for grid point extraction at an accuracy of a few mm. Furthermore, we propose a sequence of execution steps in a robotic setup, including the hand–eye calibration, which enables the direct georeferencing of multiple data sets acquired from various poses into a common coordinate system. With the proposed approach we are able to digitally reconstruct a mesh of an unknown geometry in under 10 min with an accuracy better than 5 mm. The digitally reconstructed mesh allows for computation of material needed for its construction, enabling sustainable use of concrete in digital fabrication. The accurately reconstructed digital mesh, generated based on the proposed approach in this paper, is the input for the following spraying step, allowing for generation of accurate spray trajectories.

Keywords Depth camera · Mixed pixels · Point clouds · Feedback for manufacturing · Object reconstruction · Digital fabrication · Robotic concrete spraying

1 Introduction

The use of stay-in-place formworks is one of the ways to reduce costs in construction and cut down on one-time concrete forms, which produce extensive material waste (Lloret-Fritschi et al. 2020). Stay-in-place formworks function as a structural reinforcement and may also serve aesthetic purposes (Hack et al. 2017), allowing for design freedom through free-form and complex shapes of the built structures. One application example is concrete spraying or shotcreting, where rebar meshes are used as a basis. It is a highly industrialized construction method, used primarily in modern tunneling and slope reinforcements, either in a manual or

semi-automatic mode (Fig. 1). The high-velocity application of concrete results in a highly consolidated material (Beaupre 1994). Recent advances in robotics in the area of architecture and civil engineering pushed the development of robotic spraying, where the goal is to automate the process as much as possible. Examples include spraying of plaster on flat walls (Jenny et al. 2020) or concrete onto carbon fibre or reinforcement meshes (Taha et al. 2019) (Fig. 1). In this paper we focus on the latter. To make the process automatic, the first step is to detect the overall location and shape of the mesh. This step is crucial for the spraying process and provides an input for spray trajectory generation. Accurate spray path trajectories are needed for accurate material application, which needs to be done at specific constant spraying distances, as well as to avoid collisions between the robot arm and the mesh.

Particular challenges are, e.g., handling of production errors (e.g., material fall down, insufficient layer thickness) or of unknown or imprecise geometries of work pieces (Wangler et al. 2016; Studio 2017). One of the solutions is to include adequate equipment, sensors, data processing, and feedback control in the setup, which

✉ Valens Frangez
valens.franzez@geod.baug.ethz.ch

¹ Geosensors and Engineering Geodesy, Institute of Geodesy and Photogrammetry (IGP), ETH Zürich, Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland

² Gramazio Kohler Research, Institute of Technology in Architecture (ITA), ETH Zürich, Stefano-Franscini-Platz 1, 8093 Zurich, Switzerland

Fig. 1 Manual spraying onto a rebar mesh during skate park construction, Spohn Ranch©(left). Wall reinforcement with semi-automatic spraying using a shotcrete machine robot, SIKA©(middle). Robotic concrete spraying setup (Taha et al. 2019) (right)

integrate real-time (geometric) information into the fabrication process. State-of-the-art robotic processes in digital fabrication thus include sensors, such as depth cameras, laser scanners, or stereo camera systems. However, we think their use is not yet sufficiently scrutinized (Schuler and Sawodny 2019; Wang et al. 2020).

One of the ways to digitally reconstruct thin objects, such as rebar meshes, is to scan them using laser scanners. However, the acquired data are usually influenced heavily by mixed pixels, a bias occurring when the laser beam hits surfaces at different distances simultaneously (Chaudhry et al. 2021), i.e., the rebar structure and the background in this case. Kim et al. (2020) proposed an approach of automatic rebar dimensional inspection based on terrestrial laser scanning (TLS), by implementing a noise reduction and mixed pixel removal algorithm. Furthermore, Wang et al. (2017) makes use of both 3D and color features to train a support-vector machine classifier that extracts the points corresponding to the rebars. This approach substantially relies on the information provided by RGB images, which represents a further major disadvantage and will be discussed in the next paragraph. The topic of scanning wire-objects is also under research for applications, such as inspection of power lines using mobile mapping data, see, e.g., Sánchez-Rodríguez et al. (2019). The approaches using TLS data are associated with challenging registration of data sets acquired from multiple views, occlusions, which are especially problematic in dense rebar mesh layouts, and computationally intensive, time-consuming processing of large data sets. In addition, some of the mentioned approaches rely on prior information for the reconstruction of objects, e.g., the circular shape of the wire and its diameter, meaning when the wire specifications change, the user has to adapt the processing code. When acquisition is done from a larger distance, the scanner resolution and the number of points on the wire may result in further challenges. Overall, scanning of rebar meshes is not yet a solved problem.

Another way to extract rebar meshes is using (stereo) RGB camera systems. An example of such a system was developed in the mesh mould project (Gifftthaler et al. 2017), where the position of the steel wire important for the next construction step had to be identified, such that the robot tool can clamp onto it. Damage inspection methods for steel-reinforced concrete structures are another research domain, where extraction of rebars is relevant. A few approaches exist in the literature (Xu et al. 2019; Miao et al. 2021), where areas with exposed rebars are detected using single RGB images. With these approaches it is possible to detect faulty areas by identifying the shape of the rebar wire in the image as well as using other cues of the exposed and damaged concrete structures. The major disadvantage of all RGB image-based methods is that they depend heavily on the external illumination. This significantly reduces the measurement robustness for applications in challenging environments with changing illumination conditions, such as construction sites or fabrication workshops. Furthermore, in the case of stereo systems, the need for a common field of view of the two cameras and the occlusions reduce the flexibility of the method.

In this paper, we build on top of the two mentioned ways of rebar detection: we apply image processing steps to depth images for 3D rebar mesh digitization. By this we are able to cut down on long processing times, as well as mitigate the mixed pixel effect on the final results. The starting point for the acquired depth images is an approach known in computer vision for applications, where lines have to be extracted in near real time, e.g., for road line extraction (Deng and Wu 2018) or automatic Sudoku solving with grid recognition (Kamal et al. 2015). We use commercially available time-of-flight (ToF) depth cameras to acquire the geometric data with an accuracy of a few mm over distances of up to 6 m, based on an operating principle shown in Hansard et al. (2012). The 3D observations are a result of an active illumination of the environment using a near infrared light source, making the acquisition largely independent of the external

illumination. In addition, the measurements are done from a single view point, which eliminates the common-field-of-view challenge, as is the case in the stereo systems.

Registration or alignment of point clouds, acquired at various scanner locations requires a set of transformation parameters, i.e., translations and rotations (Friedli and Wieser 2016). In a robotic setup, where sensors can be mounted on the robot arm, this challenge can be solved by calibrating the system and determining the hand–eye (HE) calibration parameters (Wang et al. 2020). Several approaches exist in the literature, where a checkerboard target is acquired by either RGB or depth cameras. Both types of cameras yield intensity in addition to 3D information (Tsai and Lenz 1989; Avetisyan et al. 2014; Wu et al. 2019). The purpose of a 2D checkerboard target is to provide a stable reference between various camera acquisitions each of which allows estimating the respective position and orientation of the camera (Albarelli et al. 2009). Furthermore, approaches of HE calibration exist, where a 3D calibration field is used, for cases when only 3D information is obtained with a depth camera (Yang et al. 2018; Kahn et al. 2014). Those are generally less accurate than the image-based solutions, because of the lower accuracy of the 3D sensors compared to the achievable accuracy with RGB cameras.

The main contributions of our paper are (i) a pipeline for accurate and efficient digital reconstruction of rebar meshes including an experimental evaluation of the accuracy, and (ii) an innovative sequence of steps for detecting the mesh in space within a robotic setup. The pipeline for rebar mesh reconstruction is depicted in Fig. 2, indicating two inputs to the process, i.e., the depth images acquired with the depth camera and the robot poses recorded by the robot arm control unit. The depth images are used for extraction of the grid points of the mesh object and then georeferenced using robot pose information complemented by the hand–eye calibration parameters. The output of the pipeline is a digitally reconstructed mesh object. Overall, our proposed approach offers a fast, autonomous, accurate, and robust solution, which can be used for detecting and reconstructing complex double-curved rebar structures or other structures composed of linear elements. Further contributions are the demonstration of

an innovative application of machine vision and point cloud data processing to digital fabrication, and the demonstration of the potential of state-of-the-art depth cameras to support robotic applications by contact-less surface digitization with mm-level accuracy.

The calibration of the robotic setup to obtain HE transformations and its evaluation are shown in Sect. 2. In Sect. 3 we assess mixed pixel influences on the acquired scans of rebar meshes. The proposed approach of depth image processing is described in detail in Sect. 4 and the accuracy assessment in Sect. 5. The sequence of steps of detecting the mesh structure within a setup is described in detail in Sect. 6.

2 System calibration

When depth cameras are rigidly mounted on a robot arm, the acquired sensor data at various robot poses can be directly georeferenced into a common coordinate system (CS). However, to be able to do this, one must determine the relation between the internal coordinate system (ICS) of the camera and the tool coordinate point (TCP) of the robot. This can be achieved by a so called HE calibration, by, e.g., making use of one of the existing methods from the literature (see Wu et al. 2019 for comparison between various methods). In this paper, we use the Tsai calibration method (Tsai and Lenz 1989), selected because of its high accuracy and computational speed (Wu et al. 2019; Marchand et al. 2005), which are relevant for the quasi-real-time application presented herein.

A checkerboard target (in our case A1 format with 65 mm grid size), placed in the scene is acquired at 20 robot poses; however, at least six are required to solve for the HE calibration transformation matrix. These poses are randomly distributed in terms of their orientation and position on a half-sphere centered around the center of the target. At each pose, a data set is acquired consisting of a recorded TCP position and orientation in the robot base coordinate system (RBCS), i.e., $\mathbf{M}_{TCP}^{RBCS}$, and an intensity image of the checkerboard target used for estimation of $\mathbf{M}_{ICS_j}^{Target}$, where j is the camera index (see Fig. 4, showing only two of the four cameras). An example pose during the HE calibration at our setup is shown in Fig. 3. The robot used is an ABB IRB 60/2.05 which has a specified pose repeatability of 0.06 mm and pose accuracy range of 0.5 mm to 1 mm (ABB Robotics 2021). This setup is used throughout the whole paper. We estimated the transformation matrix $\mathbf{M}_{ICS_j}^{Target}$ based on modeling the extrinsic camera orientation using the OpenCV Python library (Bradski 2000). The redundant observations improve the solution and allow for elimination of gross errors. Having this information, one can solve for the three unknowns of the rotation matrix $\mathbf{R}$ and the three unknowns

Fig. 2 Acquisition and processing pipeline overview

Fig. 3 Image showing one of the poses during the HE calibration procedure

Fig. 4 Diagram of various CSs within the robotic setup relevant for the HE calibration procedure. Only two out of four cameras are visible on this diagram

of the translation vector $\mathbf{t}$ contained in the $\mathbf{M}_{\text{TCP}}^{\text{ICS}_j}$, using the relation

$$\mathbf{M}_{\text{ICS}_j}^{\text{Target}} \cdot \mathbf{M}_{\text{TCP}}^{\text{ICS}_j} = \mathbf{M}_{\text{RBCS}}^{\text{Target}} \cdot \mathbf{M}_{\text{TCP}}^{\text{RBCS}} \quad (1)$$

Now, all acquired data sets with the camera can be expressed in the RBCS using

$$\mathbf{b}_i = \mathbf{R}_{\text{TCP}}^{\text{ICS}_j} \cdot \mathbf{p}_i + \mathbf{t}_{\text{TCP}}^{\text{ICS}_j} \quad (2)$$

where $\mathbf{p}_i$ is a point in the point cloud acquired by the camera and $\mathbf{b}_i$ a point expressed in RBCS.

The standard deviations σ are computed for each of the components of the transformation matrix and each of the repeated calibrations. The values for one of these

calibrations are about 0.3 mm to 0.4 mm for the translations and 0.6 mrad to 1.5 mrad for the rotations, with the Y-axis value (vertical component of the camera's coordinate system) being the most accurate one. The same σ values were obtained for all our calibrations. This high repeatability of the accuracy is expected, since the same set of robot poses was used for the different data acquisitions and the robot setup remained stable. The values are sufficiently low for the proposed application in this paper, especially acknowledging the accuracy of the robotic arm used in the setup, as well as the depth camera itself.

To evaluate the stability of the transformation parameters $\mathbf{M}_{\text{TCP}}^{\text{ICS}_j}$, we repeated the calibration procedure several times over a day, and then again 3 days later. The rotation differences $\Delta \mathbf{v}$ and translation differences $\Delta \mathbf{t}$ are small numeric variations between the repeated calibrations and they are all far smaller than 3σ , i.e., ± 4.5 mrad for rotations and ± 1.2 mm for translations, and thus statistically insignificant. Based on the outcomes of this experiment, the calibration was sufficiently accurate and stable within the calibration uncertainty, however, we recommend carrying out the calibration once per week. This supports quality control and does not cause a significant delay of the construction process.

To empirically evaluate the quality of the implemented HE calibration a scene with three spheres, two with a diameter of 15 cm and one of 12 cm (see Fig. 5a), was scanned using a depth camera from six different robot poses at distances of about 60 cm. The data sets were directly georeferenced using the recorded robot poses and manually cleaned from mixed-pixels and background. The resulting point clouds are shown in Fig. 5b. The orthogonal distances of the measured points from best-fit spheres (diameter fixed to the known true value) have an root mean square error (RMSE) of 3.7 mm. This value represents the combined effect of measurement noise, pose uncertainties and calibration uncertainties and indicates the accuracy of surface point coordinates determined using the present system. The mm-level accuracy corresponds to the expectation based on the specifications of the depth camera and the evaluation of

Fig. 5 Image of the scene with three spheres that were scanned from different camera poses (left) and the corresponding georeferenced point clouds (right), with different colors indicating different underlying poses (right)

Table 1 Selected specifications of the Helios Lucid ToF depth camera (Lucid Vision Labs 2021); accuracy defined as measurement difference to the ground truth value, and precision as measurement repeatability

Field of view	59° x 45°
Resolution	640 × 480 pixels; 0.3 MP
Range	0.3 m – 6.0 m
Accuracy	±5 mm for up to 1.5 m
Precision	0.5 mm (< 0.5 m); 2.2 mm (< 1.5 m)

the HE calibration, and it is sufficient for the purpose of the application presented herein.

The calibration is an automatic process without the need for any user interaction, except placing the target in the scene. Including robot movements to defined poses, depth camera data acquisition, and data processing, it takes approximately 5 min to complete, out of which only about 30 s are needed for computing the HE parameters.

3 Mixed pixel effect

The most relevant specifications of the depth camera used in this paper, i.e., Helios Lucid (Lucid Vision Labs 2021) for the presented investigation are given in Table 1. Tips on its use, e.g., on suggested warm-up time to achieve the highest specified accuracy are given in Frangez et al. (2020). Practical limitations of the depth cameras are related to the changes within the operating environment (i.e., mostly temperature), mixed pixels, and multipath effects. The latter occur when light bouncing off several surfaces arrives at the same pixel of the camera. The magnitude of the effect depends on the surface conditions and the geometry of the scene (Horaud et al. 2016). Most of these errors are shown in Carfagni et al. (2017) and can be mitigated by either calibration or careful setup design, when allowed by the application. The focus of the following part of this section is on the mixed pixels, a measurement bias briefly explained in Sect. 1, and on how they affect rebar edges in point clouds.

The camera poses, in particular the viewing angles and the distance from the objects to be measured, need to be carefully planned in order to assure that there are enough non-mixed pixels. This means taking into account the dimensions of the structures, i.e., rebar diameters, and the footprint, i.e., the area in the scene that corresponds to a single pixel. Figure 6 shows the footprint size as a function of distance, and the corresponding number of pixels per 4 mm (thickness of the rebars in our case) for the camera used herein. The criteria for this selection are threefold: (i) the mesh holes need to be small enough, so the material does not fly through, (ii) the rebars need to be thick enough to provide enough area for the material to stick to

Fig. 6 Theoretical relation between the footprint size and the scanning distance (red curve) and the number of pixels falling on a single rebar (blue curves). The computation is done for a rebar with a diameter of 4 mm

the structure, and (iii) the mesh needs to be stiff enough and provide sufficient structural support, without significant deformation when material is applied to it. In general, the thicker the rebar mesh, the better it can be scanned from a given distance.

There is no configuration between the camera and a rebar mesh which would yield an image, where each pixel corresponds either fully to an area on the rebars or to an area on the background. Most pixels along the edges of the rebar will be mixed pixels representing both rebar and background. The depth associated with these pixels is subject to mixed pixel errors (Wang et al. 2017; Chaudhry et al. 2021) and represents neither the distance of the rebar nor of the background. The data processing needs to take this into account by appropriate filtering. Naturally, the goal is to have footprints much smaller than the rebar cross section such that there is a sufficient number of non-mixed pixels representing the rebar in the point cloud. We conclude from Fig. 6 that distances below about 1 m are needed for scanning the rebar meshes in our application case.

We investigated the mixed pixel effects of scans of rebar meshes, in particular the impact of the background distance to the mesh, based on preliminary observations that mixed pixels do not appear if the background is far away from the rebar. The intensity of the received signal is inversely proportional to the distance squared. This means that the return intensity of the part of the footprint at the background is much smaller than of the part corresponding to the rebar if the background is sufficiently far away. Since the signal intensities also depend on the surface reflectivity and the angle of incidence, the minimum distance of the background for negligible mixed pixel effects also depends on the material and surface finish of the rebars and background. The situation is most critical if the background is orthogonal to the lines-of-sight from the camera, and if it reflects much stronger than the rebar. Theoretically, keeping the mesh at a distance of 0.4 m from the camera and the background at distances between 0.5 m and 1.3 m behind the mesh, the ratio of the

received power from foreground (rebar) and background is between approximately 1.5 and 10 growing quadratically with distance.

In relation to this, we conducted an experiment to empirically determine the background distance beyond which the mixed pixel effect is negligible. We used a strongly reflecting background approximately perpendicular to the viewing direction of the camera and at various distances from the mesh. The depth camera was located at a fixed distance of 0.4 m from the rebar mesh, while the background was varied from approximately 0.4 m to 3 m behind the mesh. An example of a point cloud acquired with a background distance from the mesh of 0.6 m and another one acquired with a background distance of 2 m are shown from the side view as well as zoomed-in from the front in Fig. 7. The grid points were extracted from these point clouds according to the approach explained in the following Sect. 4. Since the scanned mesh was quasi-planar, the accuracy of the extracted points could be evaluated by computing the RMSE of the plane fit. In the case when the background was closer to the mesh (1a, 1b), the extracted grid points are significantly affected by mixed pixel effects, the extracted rebar surface is concave instead of planar, and the RMSE at the rebar surface is about 30 mm. In the case of the background at a distance of 2 m the point cloud is much less affected by mixed pixels, and the RMSE of the extracted rebar surface is 6 mm. Overall, mixed pixels do not affect the rebar extraction critically any more with background distances from the mesh beyond 1.8 m in our experiments. While this value only holds for the specific camera, rebar, rebar distance, and background used in this experiment, the more general and thus important conclusion is that the setup planning and design for the robotic spraying should allow for sufficiently large background distance.

Fig. 7 Two point clouds used for grid point extraction and mixed pixels analysis (see text). Parts of side view (1a, 2a) and front view (1b, 2b) plotted with different scale for scans with rebar at 0.4 m and background at 0.6 m (1a, 1b) and 2 m from the mesh. The two point clouds shown here are cutoff at the approximate distance of 0.2 m

4 Grid point extraction approach

A suitable background distance during acquisition significantly mitigates the amount of mixed pixels as well as their influence on the point cloud. However, mixed pixels are still present in the data set and the extraction of the rebar from the raw point clouds needs to take this into account. We propose an image-based approach combining established image processing tools applied to the depth images. The first part of the proposed solution is the extraction of grid points from the raw data, i.e., points representing the (nearly orthogonal) intersection of rebar elements. We carry out the grid point extraction using the following five steps: (i) image pre-processing, (ii) horizontal and vertical line area extraction, (iii) intersection of lines to determine grid points, (iv) assigning depth to the grid point from the initial depth image, and (v) image post-processing. Their details are as follows:

- (i) Image pre-processing: The depth image of the acquired scene is first normalized to a range of [0, 255] (see the scene that was acquired in Fig. 8a), so the image

Fig. 8 Example for rebar grid point extraction for on a single depth data set: **a** an image of the mesh used in the experiment, **b** depth image, **c** extracted horizontal and **d** vertical areas, **e** conjunctions, **f** and depth image with indicated intersections using black dots as a result of the extraction algorithm

processing operations can be carried out using the OpenCV Python library (Bradski 2000). The image is then upsampled to double its original size using a bilinear interpolation (Fig. 8b). Although upsampling does not improve the information content of the raw data, we found that it improves the results obtained from the subsequent image processing steps, in particular in step (ii). The choice of the interpolation method did not affect the final results strongly, but the bilinear interpolation lead to the lowest RMSE of the extracted grid points as compared to the other methods which we had tried, i.e., nearest-neighbour and cubic interpolation. The approach presented in this section performs best if the grid is aligned with the rows and columns of the image to within $\pm 10^\circ$. To assure this, the orientation is estimated by converting the binary image of the original image to a skeleton shape, i.e., a thin version of the original shape that is equidistant to its original boundaries, using an implemented scikit-image algorithm *skeletonize* (van der Walt et al. 2014). This input is then used for line search using a Hough transform (Illingworth and Kittler 1988). The deviations of the resulting angles of the lines from 0° and 90° are then used to calculate the rotation angle for approximately aligning the image of the rebar mesh with the rows and columns of the image. Finally, the image is smoothed with a Gaussian filter and morphologically closed using dilation followed by an erosion (Comer and Delp 1999). The goal of closing is to remove small holes and disconnected components in the image while preserving the shape of the grid. The image shown for the processing step depiction in Fig. 8 was acquired at a distance of 0.45 m and with a background distance of 2 m. 30 frames were temporally averaged into a single image, with each frame acquired with an exposure time of 250 μ s.

(ii) Horizontal and vertical line area extraction: To extract the horizontal line areas, the edges of the horizontal mesh structure of the image have to be detected. This is done by using a Sobel operator that outputs the first order derivative in the horizontal direction. The image is then binarized and the line areas are expanded using a morphological dilation. The image at this step is shown in Fig. 8c. The same procedure is repeated for the vertical lines (see Fig. 8d).

(iii) Determination of grid points: Once vertical and horizontal lines are identified, a logical conjunction between the two images is carried out, resulting in pixel blobs indicating grid areas (Fig. 8e). Blob contours are then computed, followed by pixel centroid computation for each of the blobs.

(iv) Assigning depth to the grid point: The output pixels at this step indicate the locations of the grid points in the image (Fig. 8f). We can now obtain the set of corresponding grid points in 3D space by extracting depth

values in the original depth image at the identified grid point locations.

(v) Image post-processing: Finally, the 3D points are filtered to eliminate grid points which are implausible and thus likely erroneous. This can be done using prior knowledge based on the assumption that the actual rebar mesh corresponds to the planed one within certain tolerance limits. In our case this is particularly easy, because our experiments involve a regular rebar mesh with fixed grid size. Therefore, we carry out the plausibility control for each grid point by searching for its four nearest neighbours and comparing their respective distance to the planned rebar mesh grid size known from the design. Points whose distances agree with the expected ones to within 3 mm are kept, the rest are eliminated from the grid point set. The extracted grid points correspond to the front surface of the rebar mesh. If, e.g., a grid point representing the center of the rebar or the center of the weld between the longitudinal and transverse rebar is needed, an offset would have to be added at this step. The direction of the offset can be calculated from the nearby grid points, and the magnitude of the offset can be taken from the design of the mesh, if needed.

The presented algorithm herein is robust to be used for scanning of wide variety of conventional construction rebar meshes, which form a grid-like structure and have quasi-perpendicular intersections between the transverse and longitudinal rebars. The thicker the rebars, the better for the scanning and corresponding data processing algorithm. However, only up to a certain limit, if they are too thick, the morphological operators may not work any more, in particular near the borders of the image. The algorithm might also be suitable for grid point extraction of robotically fabricated structures, such as, e.g., mesh mould (Dörfler et al. 2019); however, this would require further investigation and possible adaptation of the algorithm. Here, instead of long transverse rebars, shorter rebar pieces are used to connect the longitudinal rebars. The investigation of the grid point extraction of the meshes which are significantly non-planar or when several of them are overlapping is left for future work.

5 Grid point accuracy assessment

Different acquisition distances, viewing angles, as well as rebar diameters, influence the final result of the extracted grid point differently. It is assumed, the 3D accuracy of the extracted grid points of the scanned rebar meshed can be on the level of the specified depth camera accuracy (Table 1). To assess this, we conducted an experimental analysis and acquired a set of point clouds at various

Fig. 9 Data acquisition configurations for accuracy assessment of grid point extraction at various distances and angles of the camera with respect to the rebar mesh (left). Sketch showing ICS and ECS on the camera front face, with the latter one realized with the T-probe measurements on the three front-face screws (right)

distances and angles of the camera with respect to the rebar mesh structure. Figure 9 (left) shows the configurations at which acquisitions were done with distances from 0.4 m to 1.4 m and angles from 0° to 50°. A laser tracker (LT) Leica AT960, a Leica LAS handheld triangulation scanner, and a Leica T-probe were used for acquiring the ground truth data with an accuracy on the level of 60 μm .

To evaluate the accuracy of the depth camera, the position and orientation of its ICS had to be determined. For this, a temporary LT coordinate system (LTCS) was established in the lab using three spherical mounted reflectors. The external camera coordinate system (ECS) is realized with the T-probe measurements (Fig. 9, right) using three front-face screws. Then, to establish a relation between ICS and ECS, at least six pairs of data sets had to be acquired to solve for six unknowns of the transformation matrix (i.e., three rotations and three translations), each one at a different camera pose with respect to the target. A single data pair corresponds to an intensity image of a checkerboard target acquired with a depth camera with data expressed in ICS, and the ESC acquired with a T-probe. This data set is then used to compute the transformation between ICS and ECS.

We do this again according to the Tsai HE calibration method (Tsai and Lenz 1989), which is explained in Sect. 2. Here, instead of the RBCS we use an arbitrary coordinate system LTCS established with the LT, and instead of the TCP we use the ECS, which is established with the T-probe. The standard deviations of the estimated translations and orientations of the transformation matrix $\mathbf{M}_{\text{ICS}}^{\text{ECS}}$, are on the level of 0.3 mm and 0.7 mrad, respectively.

In the next step, all the rebar mesh data sets acquired with the camera and those done with the T-scan, are transformed into a common coordinate system. This is achieved based on the known ECS for each of the camera positions, measured with the T-probe. A point $\mathbf{s}_i$ measured with the T-scan, can be expressed in ICS as $\mathbf{q}_i$ with

$$\mathbf{q}_i = \mathbf{R}_{\text{ICS}}^{\text{LTCS}} \cdot \mathbf{s}_i + \mathbf{t}_{\text{ICS}}^{\text{LTCS}}, \quad (3)$$

where $\mathbf{R}$ denotes the rotation matrix and $\mathbf{t}$ the translation vector, describing the transformation from the ICS to the LTCS.

Once the data is expressed in a common coordinate system the comparison can be done between corresponding sets of points acquired with the depth camera and the T-scan. Before comparison, the grid points within the depth camera data set are extracted with the algorithm presented in Sect. 4 to detect grid points. The ground truth data set acquired with the T-scan is used for manually selecting the grid areas, which are then used to compute the 3D centroids, corresponding to the grid point locations. Then, coordinate differences are compared for each of the three directions, which are then used to compute empirical standard deviations of differences σ , shown in Fig. 10 for a rebar mesh of diameter 4 mm. For all three directions (see axes directions in Fig. 9), σ values are on the level of 5 mm for up to the distance of 1 m. Throughout this analysis, we assume the bias is ≈ 0 . The σ values increase significantly if the distance grows beyond 1 m. The σ_z values are larger compared to the σ_x and σ_y , because σ_z are directly affected by the accuracy of the ToF measurement principle, while the latter two largely depend on the resolution of the image sensor. There is an impact of the angle on the σ values; however, it appears it is not significant for angles below 40°. The influence of the

Fig. 10 Empirical standard deviations of differences σ_x , σ_y , and σ_z ($1 \cdot \sigma$) of the grid points measured with a depth camera with respect to the LT data set, chosen as ground truth. Different colors of the curves indicate different mesh rotation with respect to the camera, as indicated with the colorbar

angle appears to be larger on σ_X and σ_Y , likely because of the way the rebar mesh structure is built. The longitudinal rebars are positioned in the front and the transverse in the back of the mesh. The rotation of the mesh about the Y-axis and the angle under which the mesh is measured with the camera results in a observation corresponding to the side of the rebar rather than from the front, therefore, causing a shift in the X- while not influencing the Y-direction.

In general, combining the accuracy and the precision of the depth camera (Table 1), the uncertainty of the LT system, and the way measurements were conducted, the results show the expected accuracy. A second data set was acquired on a different day using the same camera for the same measurement configurations. The σ values of the second data set confirmed the results presented herein. Specifically, the empirical standard deviation σ_{3D} of the values between the two data sets is on the level of 0.8 mm, which is within the range of the specified repeatability of the depth camera.

An important indicator of the success of the grid point detection is as well the number of grid points that get detected at certain distances and angles. In the data sets up to the distance of 1 m all of the grid points were detected, i.e., 156 points for the views far enough to see the whole mesh grid. For acquisition at the distance 1.4 m the number decreased to approximately 70 points. Data sets beyond the distance of 1.4 m were acquired as well; however, the number of pixels on the mesh structure was sparse (i.e., below 30) and did not allow for deriving reasonable conclusions; therefore, this part of the analysis was not included.

6 Mesh detection and digital reconstruction process

The RBCS is chosen as the reference CS within our robotic fabrication setup, meaning all of the acquired data with the depth cameras, process planning, and trajectory generation are expressed in that system. A sequence of steps was developed to most efficiently reconstruct rebar structures from several points of view, wherever in the work-space of the robot they are situated. The process is carried out in the following sequence: (i) coarse scanning, (ii) coarse reconstruction, (iii) refined scanning, (iv) refined reconstruction, and (v) spray path trajectory generation. A detailed description of each of the steps is as follows:

(i) Coarse scanning: The work-space area in front of the robot is scanned from a fixed number poses at different angles (Fig. 11a). For our particular setup the acquisition is done from 16 defined poses. The robot poses are chosen such that they do not reach too far into the work-space, therefore, avoiding possible collision with the structure itself. The mesh is always positioned in the

centre of the fabrication platform, i.e., within the optimal working range of the robot; therefore, this results in the coarse scanning positions within the range from ~0.8 m to ~1.6 m. Some of the grid points may not be detected, but this will be corrected in step (iii) with refined scanning. The output of this step is a set of depth images with their corresponding camera poses.

(ii) Coarse reconstruction: Each of the images is processed according to the grid point extraction algorithm from Sec. 4. The extracted sets of 3D coordinates of the grid points are then transformed into the RBCS using the camera pose information. Because of the acquisition distances, the detected grid points at this step do not have the highest achievable accuracy; therefore, the outcome of this process is a rough digital surface, because of these non-optimal scanning configuration (Fig. 11b). This digital surface is generated based on Open3D library functions (Zhou et al. 2018), using Poisson surface reconstruction producing a triangle mesh based on the extracted grid points.

(iii) Refined scanning: The digital surface from the previous step is used as an input to the second step of scanning, where a new set of poses is automatically generated (Fig. 11c). Here, each camera view covers a part of the rebar mesh at an acquisition distance of approximately 40 cm and with an overlap between views of 30%, with both parameters chosen empirically. The output of this step is a set of depth images with the corresponding robot poses.

(iv) Refined reconstruction: The grid points are extracted from the acquired images and then transformed to the RBCS (Fig. 11d). Some of the grid points will be detected several times because of the overlaps between different views. A geometric median, i.e., a point in Euclidean space minimizing the sum of distances to the input set of points, is computed for all the extracted grid points within a neighbourhood of 5 mm, a value determined empirically, thus resulting to single extracted grid point for each mesh grid intersection. Median computation is chosen, because it is a more robust statistic to outliers than, e.g., mean. In addition, all of the points which do not correspond to any actual grid point have to be filtered out at this step. The filtering is done in a way that for each point, a condition of having a least two neighbours within a certain range has to be passed, otherwise the point is excluded from the data set. In the next step, an accurate digital surface is reconstructed out of the grid points (Fig. 11e) using Poisson surface reconstruction.

(v) Spray path trajectory generation: The digital surface of the rebar structure is an input to generate the spray path trajectory. Poses are generated at a distance of approximately 20 cm from the structure, following the local normal orientation. This distance is much larger

Fig. 11 Sequence of steps for the on-site rebar mesh scanning and reconstruction process. **a** Coarse scanning with fixed set of robot poses. **b** Coarse reconstruction. **c** Refined scanning using uniquely generated poses based on the rough surface input. **d** Grid points of the

stitched individual data sets based on the recorded robot poses. **e** Fine reconstruction. **f** Robot spray trajectory generation based on the fine surface input

than the uncertainty of the preceding surface detection ($\sigma_{3D} \approx 5$ mm) and thus collisions between the robot and the rebar mesh during the spraying process are avoided.

The digitally reconstructed surface of the rebar mesh can then be used to estimate the area, which is a relevant parameter in the following process planning. Based on this and the anticipated target structure thickness, the approximate amount of material (i.e., concrete, fibers), which is required for the construction can be computed and prepared. In a controlled production facility, this would allow for a very targeted material use with reduced amount of waste at the end of the production. In this way the process can be designed more sustainably, using an additional benefit of the advanced sensing technologies.

The whole process presented in this section was applied in an experimental setup representing a realistic application example. A mesh of approximately 2.5 m^2 was scanned using one Helios Lucid depth camera mounted on the ABB IRB 60/2.05 robot. Except the preceding calibration (see Sect. 2), the whole process took about 10 min, with the computations requiring approximately

2 min, using a computer with 15.8 GB RAM and Intel Core i5-8279U CPU with 2.4 GHz. A robot pose during the scanning of the mesh is shown in Fig. 12.

Fig. 12 Image showing one of the poses during the rebar mesh scanning procedure. Here only two of the four depth cameras were in use

7 Conclusion and outlook

In this paper we demonstrated the use of depth cameras for rebar mesh scanning and digital reconstruction for robotic spraying applications. The achievable 3D accuracy, i.e., the difference of the camera measurements to the ground truth data of the extracted grid points at distances closer than 1 m was on the level of 5 mm or better. Furthermore, we proposed an innovative step sequence for detecting and reconstruction of the unknown geometry of the rebar mesh structure. It is an autonomous process demonstrated on a sample rebar mesh of 2.5 m², which took about 10 min using one depth camera. The data collection and processing time will likely increase quadratically with shorter distance as compared to the application example herein, because the mesh area covered by the camera with one image decreases quadratically and consequently more images are needed. To be able to directly georeference the data sets acquired at various robot poses, we made use of the state-of-the-art hand–eye calibration procedure, where calibration parameters can be estimated based on the acquired intensity images and corresponding robot poses. In particular, the outputs of the presented algorithm can be used for estimation of the approximate amount of material needed for construction. This allows for reduction of material waste and increase of the process sustainability. Finally, the inclusion of sensing solutions in robotic setups for construction will open up commercial opportunities, e.g., automatic customized mass production of fabricated objects due to sufficient control during fabrication.

The approach of depth image acquisition and processing is primarily designed for a pre-fabrication facility, i.e., a controlled environment, where most importantly neither the camera nor the mesh object are exposed to the direct sunlight. This is to avoid the saturation of the image sensor, which would reduce the quality of the acquired images and, therefore, the extracted mesh grid points. In this respect, the algorithm would also work on-site in an uncontrolled environment, as long as the camera and the mesh object are protected from direct sunlight. Since the camera has its own illumination source, changing light conditions are not a problem. In addition, because the time of data acquisition is short, i.e., only a few minutes, it is unlikely that the camera temperature would change sufficiently to cause significant deviations in the depth images. If these deviations would become an issue, a correction model taking the camera temperature as an input could likely be developed to mitigate the effect.

The presented algorithm herein is robust to be used for scanning of wide variety of conventional construction rebar meshes, if they form a grid-like structure and have quasi-perpendicular intersections between the transverse

and longitudinal rebars. Even though the algorithm in this paper was demonstrated only on rebars of 4 mm diameter it is applicable also to thicker or thinner rebars, and similar accuracy as found herein is to be expected. Thicker rebar meshes are more suitable for scanning and corresponding data processing. If the rebar is thinner, shorter acquisition distances are needed for the accurate reconstruction.

The proposed algorithm might also be suitable for grid point extraction from robotically fabricated structures, such as, e.g., mesh mould; however, this would have to be further investigated and could possibly require algorithm adaptation. As a part of the future work we will assess the limits of the presented algorithm in terms of scanning significantly non-planar meshes and we will address how to scan overlapping areas between several rebar mesh matts.

Acknowledgements This research was supported by the NCCR Digital Fabrication, funded by the Swiss National Science Foundation (SNSF): NCCR Digital Fabrication (#51NF40–182887) and by Innosuisse (#35019.1 IP-ENG). The authors would like to thank Nicolas Feihl for support with the robot control.

Funding Open Access funding provided by ETH Zurich.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Albarelli A, Rodolà E, Torsello A (2009) Robust camera calibration using inaccurate targets. *IEEE Trans Pattern Anal Mach Intell* 31:376–383
- Avetisyan R, Willert M, Ohl S, Staadt O (2014) Calibration of depth camera arrays. In: *Proceedings of SIGRAD 2014, Visual Computing*, June 12–13, 2014, Göteborg, Sweden, no. 106. Linköping University Electronic Press, pp 41–48
- Beaupre D (1994) Rheology of high performance shotcrete. Ph.D. dissertation. University of British Columbia
- Bradski G (2000) The OpenCV library. *Dr. Dobb's J Softw Tools Prof Programmer* 25(11):120–123
- Carfagni M, Furferi R, Governi L, Servi M, Uccheddu F, Volpe Y (2017) On the performance of the intel sr300 depth

- camera: metrological and critical characterization. *IEEE Sens J* 17(14):4508–4519
- Chaudhry S, Salido-Monzú D, Wieser A (2021) A modeling approach for predicting the resolution capability in terrestrial laser scanning. *Remote Sens* 13(4):615
- Comer ML, Delp EJ III (1999) Morphological operations for color image processing. *J Electron Imaging* 8(3):279–289
- Deng G, Wu Y (2018) Double lane line edge detection method based on constraint conditions hough transform. In: 2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES). IEEE, pp 107–110
- Dörfler K, Hack N, Sandy T, Giftthaler M, Lussi M, Walzer AN, Buchli J, Gramazio F, Kohler M (2019) Mobile robotic fabrication beyond factory conditions: case study mesh mould wall of the dfab house. *Construct Robot* 3(1):53–67
- Frangez V, Salido-Monzú D, Wieser A (2020) Depth-camera-based in-line evaluation of surface geometry and material classification for robotic spraying. In: 37th International Symposium on Automation and Robotics in Construction (ISARC 2020)(virtual). International Association for Automation and Robotics in Construction, pp 693–702
- Friedli E, Wieser A (2016) Identification of stable surfaces within point clouds for areal deformation monitoring. In: Proc. of 3rd Joint International Symposium on Deformation Monitoring (JISDM)
- Giftthaler M, Sandy T, Dörfler K, Brooks I, Buckingham M, Rey G, Kohler M, Gramazio F, Buchli J (2017) Mobile robotic fabrication at 1:1 scale: the in situ fabricator. *Construct Robot* 1(1):3–14
- Hack N, Wangler T, Mata-Falcón J, Dörfler K, Kumar N, Walzer AN, Graser K, Reiter L, Richner H, Buchli J et al (2017) Mesh mould: an on site, robotically fabricated, functional formwork. Second Conc Innov Conf (2nd CIC) 19:1–10
- Hansard M, Lee S, Choi O, Horaud RP (2012) Time-of-flight cameras: principles, methods and applications. Springer, Berlin
- Horaud R, Hansard M, Evangelidis G, Ménier C (2016) An overview of depth cameras and range scanners based on time-of-flight technologies. *Mach Vis Appl* 27(7):1005–1020
- Illingworth J, Kittler J (1988) A survey of the hough transform. *Comput Vis Graph Image Process* 44(1):87–116
- Jenny SE, Lloret-Fritsch E, Gramazio F, Kohler M (2020) Crafting plaster through continuous mobile robotic fabrication on-site. *Construct Robot* 4(3):261–271
- Kahn S, Haumann D, Willert V (2014) Hand-eye calibration with a depth camera: 2d or 3d? In: 2014 International Conference on Computer Vision Theory and Applications (VISAPP), vol 3. IEEE, pp 481–489
- Kamal S, Chawla SS, Goel N (2015) Detection of sudoku puzzle using image processing and solving by backtracking, simulated annealing and genetic algorithms: A comparative analysis. In: 2015 third international conference on image information processing (ICIIP). IEEE, pp 179–184
- Kim M-K, Thedja JPP, Wang Q (2020) Automated dimensional quality assessment for formwork and rebar of reinforced concrete components using 3d point cloud data. *Autom Construct* 112:103077
- Lloret-Fritsch E, Wangler T, Gebhard L, Mata-Falcón J, Mantellato S, Scotto F, Burger J, Szabo A, Ruffray N, Reiter L et al (2020) From smart dynamic casting to a growing family of digital casting systems. *Cement Concr Res* 134:106071
- Marchand E, Spindler F, Chaumette F (2005) Visp for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robot Automat Mag* 12(4):40–52
- Miao Z, Ji X, Okazaki T, Takahashi N (2021) Pixel-level multicategory detection of visible seismic damage of reinforced concrete components. *Comput Aid Civil Infrastruct Eng* 36(5):620–637
- ABB Robotics (2021) Product specification - IRB 4600. <https://new.abb.com/products/robotics/industrial-robots/irb-4600>. Accessed 31 Mar 2021
- Sánchez-Rodríguez A, Soilán M, Cabaleiro M, Arias P (2019) Automated inspection of railway tunnels' power line using lidar point clouds. *Remote Sens* 11(21):2567
- Schuler B, Sawodny O (2019) Spray pattern analysis using wet-mix concrete for model based process control towards automated construction. In: 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE). IEEE, pp 661–666
- Studio E (2017) Robotic fabrication techniques for material of unknown geometry. In: Humanizing Digital Reality: Design Modelling Symposium Paris 2017. Springer, p 311
- Taha N, Walzer AN, Ruangjun J, Bürgin T, Dörfler K, Lloret-Fritsch E, Gramazio F, Kohler M (2019) Robotic aerocrete—a novel robotic spraying and surface treatment technology for the production of slender reinforced concrete elements. In: Architecture in the Age of the 4th Industrial Revolution—Proceedings of the 37th eCAADe and 23rd SIGraDi Conference, vol 3. CumInCAD, pp 245–254
- Lucid Vision Labs (2019) Time-of-flight forged ahead: design tips to boost 3D performance and cut integration time and cost, user manual, v.1.3, 18.10.2019
- Tsai RY, Lenz RK et al (1989) A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. *IEEE Trans Robot Autom* 5(3):345–358
- van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T (2014) scikit-image: image processing in Python. *PeerJ*
- Wang Q, Cheng JC, Sohn H (2017) Automated estimation of reinforced precast concrete rebar positions using colored laser scan data. *Comput Aid Civil Infrastruct Eng* 32(9):787–802
- Wang Z, Fan J, Jing F, Deng S, Zheng M, Tan M (2020) An efficient calibration method of line structured light vision sensor in robotic eye-in-hand system. *IEEE Sens J* 20(11):6200–6208
- Wangler T, Lloret E, Reiter L, Hack N, Gramazio F, Kohler M, Bernhard M, Dillenburger B, Buchli J, Roussel N et al (2016) Digital concrete: opportunities and challenges. *RILEM Tech Lett* 1:67–75
- Wang Z, Zhen Z, Lin Z, Wen T, Guo C, Che H (2020) An adaptive industrial robot spraying planning and control system. In: IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, IEEE, pp 4739–4743
- Wu J, Sun Y, Wang M, Liu M (2019) Hand-eye calibration: 4-d procrustes analysis approach. *IEEE Trans Instrum Measure* 69(6):2966–2981
- Xu Y, Wei S, Bao Y, Li H (2019) Automatic seismic damage identification of reinforced concrete columns from images by a region-based deep convolutional neural network. *Struct Control Health Monit* 26(3):e2313
- Yang L, Cao Q, Lin M, Zhang H, Ma Z (2018) Robotic hand-eye calibration with depth camera: A sphere model approach. In: 2018 4th International Conference on Control, Automation and Robotics (ICCAR). IEEE, pp 104–110
- Zhou Q-Y, Park J, Koltun V (2018) Open3d: A modern library for 3d data processing. *arXiv preprint arXiv:1801.09847*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.