Geometric Feedback System for Robotic Spraying

PFG (2022) 90:531–542
https://doi.org/10.1007/s41064-022-00219-7

ORIGINAL ARTICLE

Geometric Feedback System for Robotic Spraying

Valens Frangez[1] · Nizar Taha[2] · Nicolas Feihl[2] · Ena Lloret-Fritschi[3] · Fabio Gramazio[2] · Matthias Kohler[2] · Andreas Wieser[1]

Received: 14 April 2022 / Accepted: 31 August 2022 / Published online: 18 October 2022

Abstract

In this paper, we tackle the task of replacing labor intensive and repetitive manual inspection of sprayed concrete elements with a sensor-based and automated alternative. We present a geometric feedback system that is integrated within a robotic setup and includes a set of depth cameras used for acquiring data on sprayed concrete structures, during and after fabrication. The acquired data are analyzed in terms of thickness and surface quality, with both sets of information then used within the adaptive fabrication process. The thickness evaluation is based on the comparison of the as-built state to a previous as-built state or to the design model. The surface quality evaluation is based on the local analysis of 3D geometric and intensity features. These features are used by a random forest classifier trained using data manually labelled by a skilled professional. With this approach, we are able to achieve a prediction accuracy of 87 % or better when distinguishing different surface quality types on flat specimens, and 75 % when applied in a full production setting with wet and non-planar surfaces. The presented approach is a contribution towards in-line material thickness and surface quality inspection within digital fabrication.

Keywords Feedback system · Geometric quality assessment · Depth camera · Digital fabrication · Robotic spraying

Zusammenfassung

Geometrisches Rückkopplungssystem für robotergesteuertes Betonspritzen. In diesem Beitrag schlagen wir ein berührungsloses Messsystem für die Form- und Oberflächenkontrolle während des robotergestützten Betonspritzens vor. Es soll die durchgehende Automatisierung dieses Herstellungsprozesses ermöglichen, der für das digitalisierte Bauen erforscht wird und derzeit noch wiederholte manuelle und visuelle Prüfung erfordert. Als Sensoren setzen wir vier Time-of-Flight Kameras ein, die symmetrisch um die Spritzdüse verteilt am Ende des Roboterarms montiert sind und die einzelnen Abschnitte jeweils unmittelbar nach dem Spritzen aufnehmen. Aus den georeferenzierten Tiefen- und Intensitätsbildern werden sowohl die aufgetragene Dicke als auch die Oberflächenqualität analysiert. Die Dicke leiten wir durch Vergleich mit einer früheren Aufnahme oder dem Entwurfsmodell ab. Die Oberflächenqualität beurteilen wir mit einer Random-Forest-Klassifikation ausgehend von geometrischen und radiometrischen Features. Wir erreichen eine Genauigkeit von 87 % oder besser für die Klassifikation bei vorgefertigten flachen Proben und 75 % bei Anwendung in einer realen Produktionsumgebung mit nassen und unebenen Oberflächen. Der vorgestellte Ansatz ist ein Beitrag zur Automatisierung mittels kamerabasierter Messungen im digitalisierten Bauen.

✉ Valens Frangez
frangezv@ethz.ch

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

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

[3] Fabrication- & Material-Aware Architecture, Academy of Architecture, Università della Svizzera italiana (USI), Via Alfonso Turconi 36, 6850 Mendrisio, Switzerland

1 Introduction

Recently published research in the field of digital fabrication demonstrates potential for the architecture, engineering, and construction industries, by not only investigating novel geometries, new materials, performance and functionalities, but also the fabrication processes themselves and the technological developments behind them (Eversmann et al. 2017; Wangler et al. 2019). Despite sophisticated robotic setups which are built to automate the production process as much as possible, tasks such as quality inspection during and after fabrication still heavily depend on manual work carried out by skilled professionals (Mineo et al. 2020). The inspection remains a labor intensive and repetitive activity that in many cases can be replaced by digital technologies available nowadays. The latter starts to get attention in published literature, where fabrication setups augmented with various sensing systems show potential of in-line process improvement via feedback control, as e.g. demonstrated in Wolfs (2019), Rodríguez-Gonzálvez and Guidi (2019), Bard et al. (2018). In particular, this feedback is relevant in additive manufacturing applications with cementitious materials which are unpredictable and hard to model in their dynamic transition from the soft to hard state. This unpredictability remains a complex challenge which can jeopardize the visual and structural qualities of the fabricated pieces (Wolfs et al. 2018; Sutjipto et al. 2018). Automatic quality inspection to detect and classify manufacturing errors is usually carried out by identifying discrepancies between virtual design models and their corresponding as-built structures (Buswell et al. 2020). Such way of carrying out inspections promises to bring the fabrication field one step closer towards mass customization and production (Bard et al. 2018).

Different types of sensor feedback systems exist that can support the fabrication process, however, the focus of the presented work is on systems that capture and provide visual as well as 3D information of the objects. In Bard et al. (2018), the authors used an RGB camera system and a visual feedback loop for the purpose of quality evaluation of plastered surfaces and the acquired data are analyzed using convolutional neural networks to identify manufacturing errors. Furthermore, Jenny et al. (2020) explored the design potential of bespoke plastered elements by using a depth camera capturing the 3D shape of the plastered objects. A terrestrial laser scanner (TLS) was used for evaluating the shape of concrete elements produced with shotcrete 3D printing in Maboudi et al. (2020). A 1D distance sensor mounted on an end-effector was used to capture the height of 3D concrete printed layers, information that was then used to keep the nozzle at a constant distance from the surface of the printed object (Wolfs et al. 2018). Moreover, Kazemian et al. (2019) explored not only the nozzle distance from the surface but also the width of the layer that is extruded, thus adapting the process in a way to reach the desired outcome. Off-line concrete damage inspection using TLS and involving surface 3D feature computation is presented in Hadavandsiri et al. (2019), and inspection using TLS intensity image classification in Zaczek-Peplinska and Osińska-Skotak (2018). According to the authors knowledge, the potential of geometric feedback systems is not yet fully exploited and further developments are needed, as will be presented in this contribution.

Robotic concrete spraying provides the context of the research presented herein. One of the research goals is to automate the quality inspection task during and after fabrication of concrete elements (Taha et al. 2019; Ercan Jenny et al. 2021). The evaluation of thickness as well as the surface quality of such elements is so far still predominantly done manually by skilled professionals. E.g. average thickness is determined by measuring the length of the wet-part of a pin that is inserted into the freshly sprayed concrete at multiple positions. Furthermore, the surface quality and the visual appearance are assessed by subjective judgement of a skilled professional. In this paper, we investigated how to replace both manual inspection steps, i.e. thickness and surface quality evaluations, with digital alternatives. To achieve this, we made use of 3D sensors and corresponding data processing tools, such as machine learned models as a substitute for subjective quality perception of a skilled professional. Skill transfer from a human to robots via sensory inputs is not a new research field, however, it is becoming more and more relevant for construction automation and similar applications (Liu et al. 2020). In our case, we employed a set of industrial depth cameras Helios Lucid to acquire 3D and intensity information of the object under construction. The selected sensor is suitable for the application at hand, because of its comparatively low cost, and fast and accurate acquisition, which is due to the own illumination source entirely independent of the external illumination conditions and interferences (Frangez et al. 2021a).

The main contributions of the presented study are: (i) we propose a geometric feedback system (GFS) within the robotic spraying workflow that autonomously carries out the inspection of the thickness and surface quality of elements during and after construction. The GFS provides thickness and surface quality maps that are used by the robotic fabrication process to determine the required actions for achieving the desired geometry and surface quality. (ii) We demonstrate the use of 3D and intensity information provided by a set of depth cameras for deriving the surface quality maps of sprayed concrete samples. In this regard, we propose a pipeline of local feature analysis and supervised classifier incorporating different surface features and surface types. The data for training and testing of the classifier were obtained from the manual assessment of samples by a skilled professional.

In Sect. 2, we present the GFS and its integration within the overall workflow. The surface quality assessment is then described in detail in Sect. 3. The experimental setup, including the explanation of the individual setup components and the measurement process are presented in Sect. 4. The results of the presented work are shown in Sect. 5. The main conclusions and outcomes of the study are summarized in Sect. 6.

2 Geometric Feedback System

2.1 Workflow

The GFS workflow comprises rules and conditions for thickness and surface quality evaluations, production operator and 3D depth camera system inputs, and reaction scenarios (i.e. either revision, continuation, or human intervention). There are generally four production steps in the process of concrete spraying, namely (I) spraying of the glass fiber reinforced concrete (GFRC) layer, (II) compaction of the GFRC layer, (III) spraying of concrete, and (IV) troweling (i.e. spreading and shaping) of the concrete (Taha et al. 2019) (see Fig. 1). After each step, the surface can be modified for the duration of the mix-specific open-time before the concrete starts to set. Each of the subsequent steps can only be carried out, if the produced surface in the previous step was sufficiently thick and of sufficient quality, i.e. both determined according to the design model of the fabricated piece.

The thickness check and the surface quality check are carried out for selected production steps and are, therefore, the integral parts of the GFS. The thickness check refers to evaluating whether enough material is deposited onto the surface based on the reference information obtained from the design model, and the surface quality check refers to evaluating whether the desired surface outcome is reached or not based on the surface type classification.

In the thickness check, the comparison between two models (e.g. between the design and the scanned surface) is a standard procedure which results in spatially mapped deviations (see e.g. Buswell et al. 2020 for the use in construction applications). The deviations are grouped into three categories, i.e. sufficient, too much, and too little material, based on two defined thresholds. The thresholds are directly related to the current fabrication setup and on-site conditions. The possible reactions to this automatic assessment are: (i) when the layer is sufficiently thick the process continues, (ii) when the layer is too thick the process stops and removal of material is needed, likely requiring manual intervention at the current technological development, and (iii) when the layer is too thin more material is sprayed. The measured deviations are used to adapt the speed and distance of the nozzle to the object to reach the wanted thickness. The relations between these parameters are determined experimentally. The process of thickness evaluation is carried out only after steps I and III, i.e. when additional material has been applied. These steps and the evaluation are repeated until a sufficient thickness is reached in all of the areas. The surface quality check does not aim at quantifying the surface quality but classifying it as *desired* or *undesired*. The two outcomes are defined based on a skilled-professional's evaluation. The process of preparing training and test samples is explained in the following subsection and the way the captured data are processed to extract the relevant surface features is given in Sect. 3. The surface check is done after steps I, II, and IV, and is repeated after each surface manipulation as long as the desired surface quality is not reached throughout the whole surface. This decision is based on trained classifiers, with one for each decision step, thus resulting in three *desired* and three *undesired* surface quality types, in total six. There is no surface check after step III, because only a sufficient amount of material has to be applied, which is then manipulated in step IV to achieve the desired surface quality.

2.2 Training and Test Samples

There are three groups of samples used in this investigation, i.e. manual flat, robotic flat, and robotic irregular samples, for the purpose of training, validation, and test of the random forest (RF) classifier. For training and validation, we use a set of samples produced by a skilled professional in a manual way and on a flat base. For testing, we use two robotically produced sets, first on a flat and second on an irregular base (Fig. 2). The purpose of these three types is to demonstrate

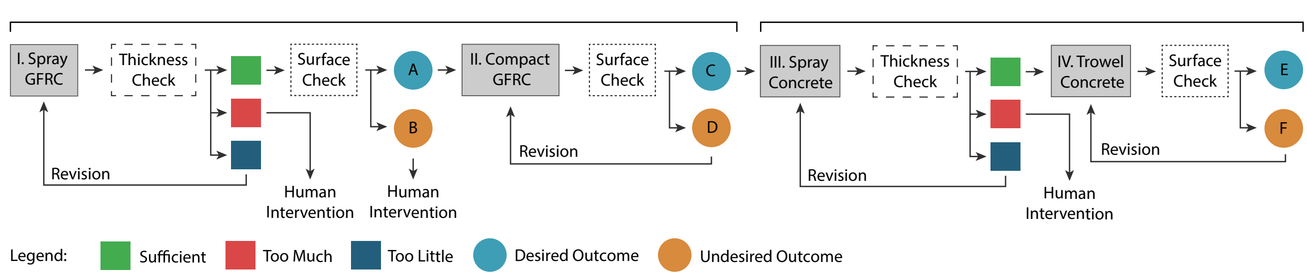

Fig. 1 Key steps within the GFS workflow process. The legend at the bottom indicates the possible outcomes, i.e. sufficient, too little, or too much thickness in the case of thickness check, and desired outcome or undesired outcome in the case of surface check. The surface types A-F correspond to the examples shown in Fig. 3

Fig. 2 Production of manual (**a**), and robotically sprayed samples (**b**, **c**). Both, **a**. and **b**. are fabricated on a flat bases, while **c**. is fabricated on an irregular one

the transferability of the surface classification developed on manual samples onto those produced by the robot, and from flat surface samples to irregular ones.

The skilled professional manually produced representative sprayed samples using a flat base, following the four production steps explained in Sect. 2.1. The professional made sure enough material is applied as well as the desired surface quality is reached. For the sake of the analysis, the professional aimed at reaching both, *desired* and *undesired* surface types, which are then labeled by the professional. This group of manually produced flat samples consisted of 20 different surface samples, i.e. more samples for each of the 6 surface quality type class than necessary, thus increasing the redundancy. The photos of samples and their corresponding depth and reflectance images are shown for representative examples of each of the surface types in Fig. 3. These surface types serve as a quality benchmark for different surface types of each production step. The same processing steps as explained for the manual samples are done using the robotic setup, first on a flat and second on an irregular mould.

3 Surface Quality Assessment

We processed the depth and intensity images to learn the RF classifier by executing the following five steps: (a) data correction, (b) label annotation, (c) feature computation, (d) supervised learning, and (e) label prediction and post-processing (see Fig. 4). In these steps, we use the manually

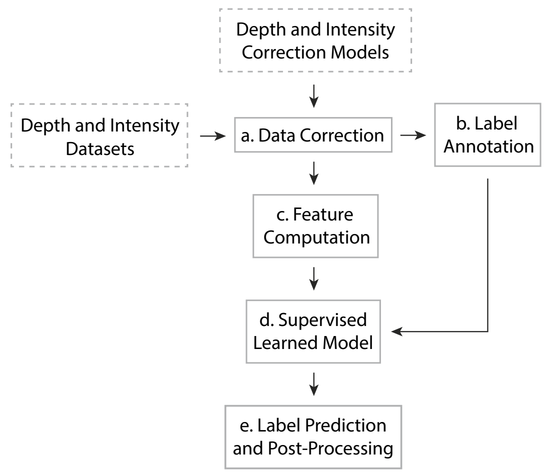

Fig. 4 Data processing pipeline overview

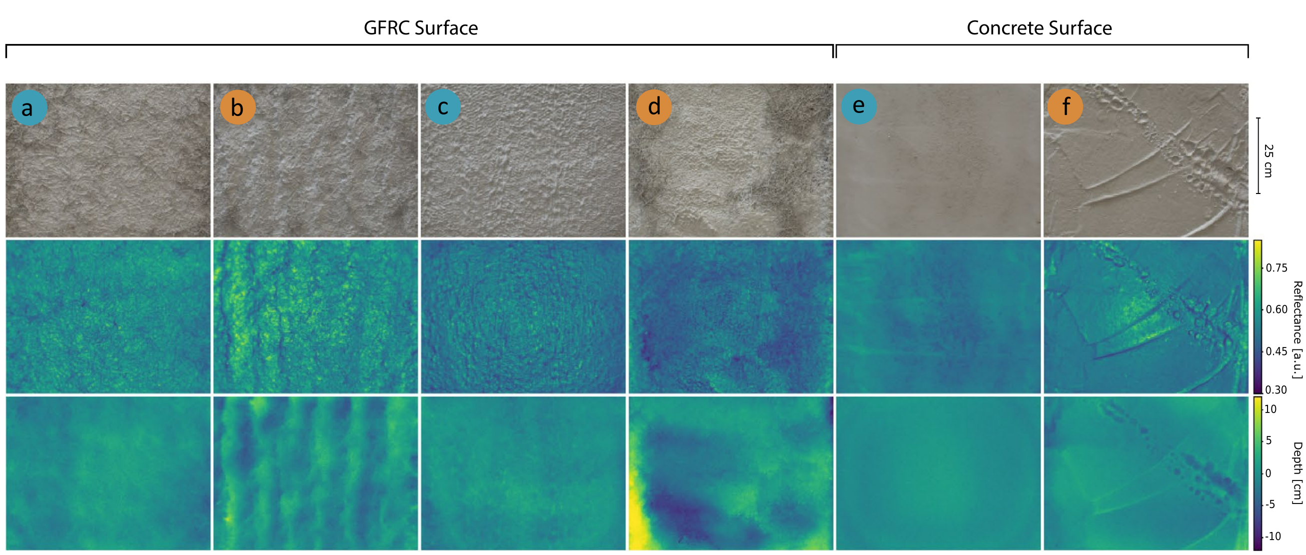

Fig. 3 Representative example images of the surface categories **A**–**F** (first row), corresponding reflectance images (second row row) and depth images (third row). The intensity and depth images shown in this figure are already corrected using the intensity and depth correction functions

sprayed flat samples as the training data. The details of the steps are as follows:

(a) Data correction: To eliminate the impacts of the scanning distance and angle of incidence as a result of measurement configurations, as well as all the camera-related inter-pixel variations, the acquired intensity images are converted into reflectance images. The latter contain only information of the reflecting surface itself (Frangez et al. 2021b). For this, we use an intensity calibration function which is determined in an independent experiment. Additionally, the depth images are corrected for all the distance-related systematics and inter-pixel related-errors, by applying an estimated error compensation function (Frangez et al. 2022). To avoid the loss of information by smoothing out the smallest surface features the images are not low-pass filtered or manipulated in any other way. The output of this step is a set of reflectance images and their corresponding depth images which are compensated for depth errors. An example dataset before and after corrections is shown in Fig. 5.

(b) Label annotation: Image labeling is an essential step in supervised learning tasks and having high-quality input data with annotations determines the quality of the learned model. We perform pixel-wise binary label annotation (Yao et al. 2016) to each of the 20 datasets, by assigning each pixel as *desired* or *undesired*. This task was carried out using the GIMP software. A result of this manual image annotation might be biased, i.e. some individual pixels, especially in the border regions between different classes, might be labeled as wrong. This so-called label noise has a negative impact on the prediction accuracy of the model, however, its effect can be slightly reduced by label filtering in the post-processing. The result of this step are manually annotated images, an example of which can be seen in Fig. 8a annotated based on Fig. 8b.

(c) Feature computation: To exploit geometric properties of the scanned surfaces, i.e. features, we perform local neighborhood analysis using reflectance and depth images. The outcome of this process are feature images, i.e. images on which surface structures are represented by a set of local feature descriptors extracted from each pixel's neighborhood. Three groups of handcrafted features are extracted from the datasets, namely geometric profile, geometric areal, and reflectance features. The profile features are handpicked from current standards in surface metrology, where mathematical surface texture is of interest (see e.g. ISO 4287 DIN EN ISO 4287:2010-07 2010). The areal features are selected from a large variety of 3D point cloud features used in similar point cloud processing applications (see e.g. Weinmann et al. 2014), and are primarily based on the eigenvalues of the covariance matrix of the local neighborhood used for computation. Lastly, the reflectance features are taken from basic image processing and feature extraction approaches (see e.g. Hassaballah et al. 2016), and are primarily aiming to exploit small-scale (i.e. sub-mm) features that cannot be captured by the geometry of the depth image, due to the accuracy limitation. The decision whether to include all the initially selected features is made after assessing the relevance and importance of each of the feature types for the prediction of the correct label, without increasing the redundancy, using the analysis of variance (ANOVA) statistical test. Eliminating irrelevant features from the computation also reduces the computational complex-

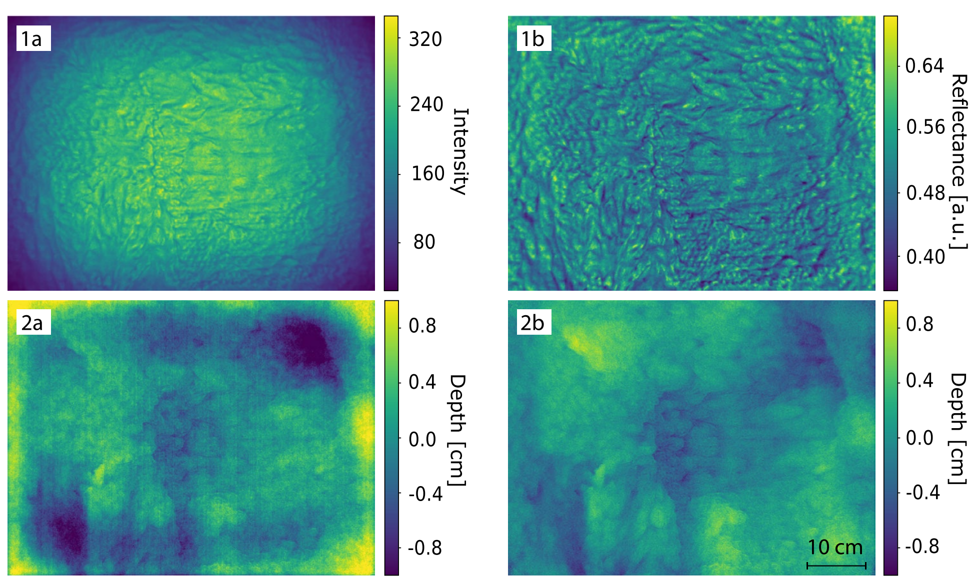

Fig. 5 Acquired intensity image (**1a**) and estimated reflectance image (**1b**). Depth image before (**2a**) and after (**2b**) correction

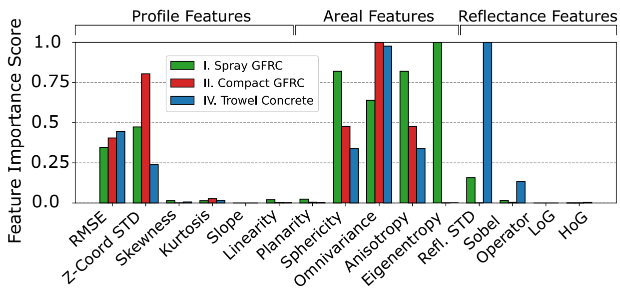

Fig. 6 Feature importance ANOVA for each of the three classifiers

ity for feature computation, making it more suitable for a real-time application. For an overview, all the features used in this analysis and their corresponding ANOVA values are shown in Fig. 6. The subset of relevant features is selected based on a threshold of approximately 0.2, computed using the median of all the ANOVA importance values, according to the default implemented in the Sklearn library (Buitinck et al. 2013). It can be observed that different types of surfaces can be exploited using different types of features and are differently important. To assess the level of mutual information captured by the three feature groups we further perform the correlation analysis between the selected subset of feature pairs computed from the 20 acquired datasets. The results can be seen in Table 1. As it can be noticed, the correlation between the 3D features, i.e. profile and areal ones (see indication in Fig. 6), is high and significantly smaller for the reflectance features, indicating the importance of reflectance information in our analysis. To describe local geometry at each image pixel (i.e. for profile and areal features), a set of points is queried from within a defined 3D search radius of the pixel of interest. This optimal radius size has to be small enough to allow detection of the finest surface details, however, big enough to average out the effects of noise present in the datasets (Hadavandsiri et al. 2019). The obtained subset of points is then used to determine the feature vector of that single point, and this procedure is repeated for all the points within each dataset. For this, we use 13 different sizes of the search radius, ranging from about 2 mm to about 200 mm in size. Each of these radii are used to extract the same set of features which are then used in the subsequent classifier analysis (see step d. for further information on the prediction accuracy) to determine which of them is able to predict the labels most accurately. Each image used in the feature extraction is prior to that padded with a number of pixels related to the search radius to avoid the errors on the border of the image (Aghdasi and Ward 1996). To extract reflectance features from the 2D reflectance images, the radius is converted to pixels, such that the corresponding kernel size could be defined. The procedure is for the rest similar as for the 3D feature computation explained above. The outcome of this step is a set of feature images.

(d) Supervised learning: After obtaining the labels for each of the pixels (step b.) and the features (step c.), we split each image dataset acquired using manually produced samples to 70 % training and 30 % validation datasets. The training dataset is used in a supervised learning process, where we learn the relationship between the feature vector of each point and the assigned label indices. The validation dataset is used to assess the prediction accuracy of the trained classifier. For this task, we try out different types of classifiers, namely nearest neighbour, support vector machine, and RF classifiers from Scikit-learn Python machine learning library (Pedregosa et al. 2011; Speiser et al. 2019). The latter performs best in terms of the prediction accuracy using the validation dataset and is most efficient in terms of the prediction computation time as well, therefore it is used in the next steps of the presented analysis. Prediction accuracy based on the validation dataset used herein is defined as the ratio of correctly labeled pixels with respect to all the pixels for each label class, averaged over the total number of classes within the dataset (Fooladgar and Kasaei 2020). We use a randomized search on hyper-parameters to determine the optimal set of them to increase the classifier performance. As an alternative approach to classification, deep learning could be considered. This may be beneficial, because it

Table 1 Correlation values between different features used in the presented analysis. The shown features are the selected subset from those shown in Fig. 6

	Z-Coord STD	Sphericity	Omnivariance	Anisotropy	Eigenentropy	Refl. STD	Sobel
RMSE	.61	.77	.82	− .77	.04	.57	.28
Z-Coord STD		.64	.66	− .64	.21	.36	.16
Sphericity			.91	− 1.00	.05	.52	.26
Omnivariance				− .91	− .02	.69	.35
Anisotropy					− .05	− .52	− .26
Eigenentropy						− .18	− .16
Refl. STD							.57

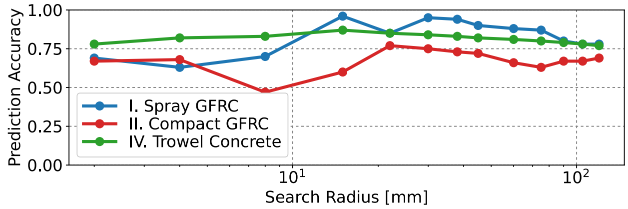

Fig. 7 Class prediction accuracy on the validation dataset with respect to the size of the search radii used for feature computation for each of the three classifiers

Table 3 Precision and recall values, as well as the percentage of FP and FN predictions for each surface type (see Fig. 3 for the surface types A–F

Type	Precision	Recall	FP	FN
A	.88	.94	.04	.02
B	.91	.85	.03	.03
C	.92	.99	.02	.01
D	.99	.96	.01	.01
E	.94	.82	.03	.03
F	.79	.90	.08	.05

allows exploiting more salient features than the handcrafted ones used herein. However, we did not have an opportunity to provide enough training data, and did not pursue this approach further, therefore. The results of the prediction accuracies for the validation dataset for each of the three steps are shown in Fig. 7. The accuracies are shown in relation to the size of the search radii. We can observe that the optimal sizes for each of the three classifiers are on the level of approximately 15 mm. Furthermore, we show prediction accuracies per label class for each of the three classifiers for the optimal search radius (see Table 2). Each label class prediction accuracy is on the level of about 80 % or better with not many false positives (FP) (i.e. pixel being labeled as *undesired outcome*, when in reality being *desired outcome*) or false negatives (FN) (i.e. pixel being labeled as *desired outcome*, when in reality being *undesired outcome*). Overall, we are able to reach 87 % of prediction accuracy using the validation dataset. We further report precision and recall values, as well as the percentage of FP and FN predictions for each surface type in Table 3.

(e) Label prediction and post-processing: The output of this step is a probability of each of the possible labels, out of which the most likely one is selected (see Fig. 8c, d). The spatial distribution of estimated labels usually exhibits a certain level of variation and is not smooth, e.g. a different label within a neighbourhood of many same labels is likely wrong. To improve the results, the label images can be filtered/smoothed using different filters that perform similarly well (Schindler 2012). For this task, we select the bilateral filter, i.e. a non-linear filter, where each pixel within the used filter kernel influences the central pixel by not only its spatial distance to it, but also by the difference of the labels within the kernel (Tomasi and Manduchi 1998). This results in stronger filtering within homogeneous neighborhoods and weaker one when a lot of discontinuities occur. The filtering improves the results in terms of smoothing the boundaries, however, also in terms of prediction accuracy from 3 % to 5 % for each of the three classifiers. The labeled images are then used to filter out patches that are smaller than the footprint area of the spray, trowel, or roller tools used in fabrication. This is done by searching the image for contours, computing their corresponding enclosed areas, and then filtering out those that are below the threshold. The purpose of this is to smooth out the small variations that are not possible to correct with the particular setup and tools used (see Fig. 8e), however, this decision depends on the operator of a particular setup. The comparison between the finally predicted labels and the judgement of a skilled professional is shown in Fig. 8f, for true, FN, and FP predictions. The outcome of this particular step are surface quality maps, which are then used in the GFS process.

Table 2 Confusion matrices for each of the three classifiers. The values per row might not add up to exactly 1, because of rounding. The surface types A–F correspond to the examples shown in Fig. 3

Spray GFRC			Compact GFRC			Trowel Concr.		
	A	B		C	D		E	F
A	.99	.00	C	.79	.21	E	.92	.07
B	.04	.96	D	.14	.85	F	.12	.87

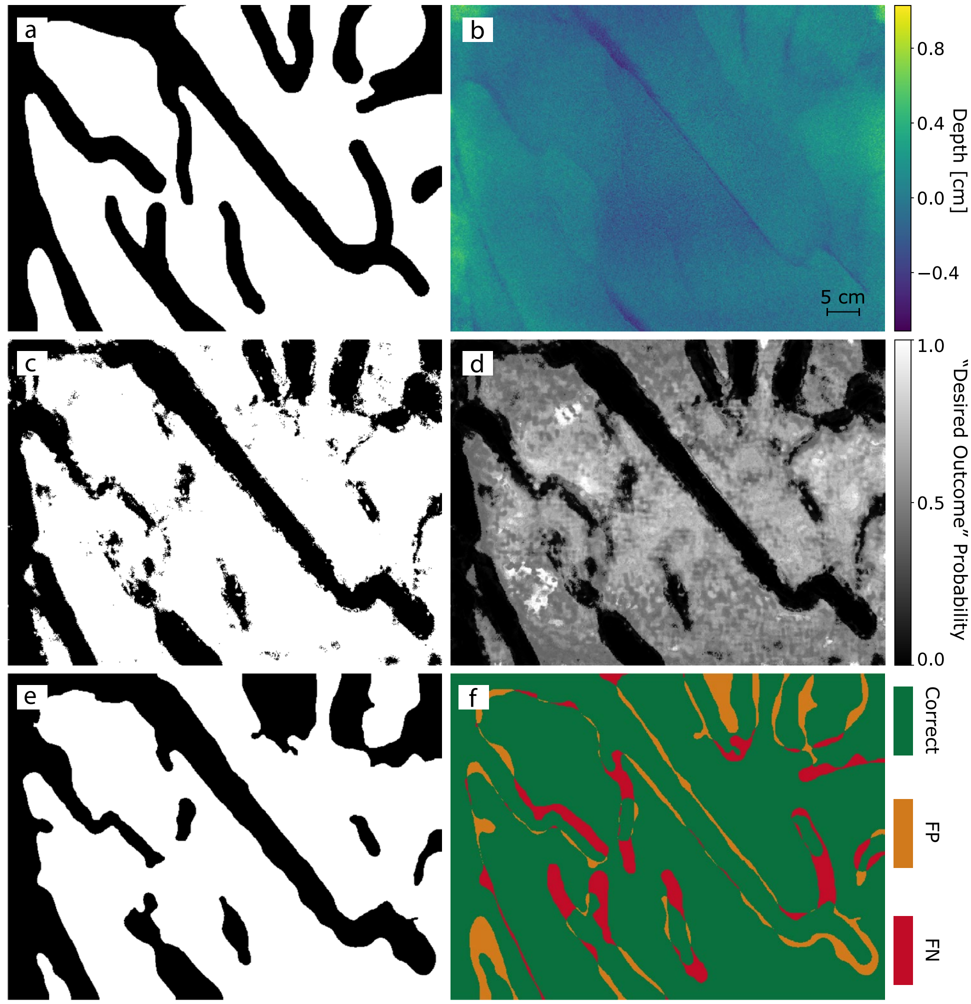

Fig. 8 Manually annotated labels (**a**), depth image of the surface sample used for annotation (**b**), estimated labels based on the model prediction (**c**), probability values for class *desired outcome* (**d**), filtered labels (**e**), and correctness map, indicating true, false positive (FP), and false negative (FN) prediction of labels (**f**)

4 Experimental Setup

4.1 Setup Components

There are two main components of the fabrication system, namely (i) a robotic arm, with a digitally controlled concrete spraying nozzle (CSN), and (ii) a set of depth cameras as a part of the GFS located on the robot end-effector. The CSN component is responsible for the spraying of concrete and GFRC. The task of the GFS component is to provide 3D information of the structure under fabrication by 3D imaging of the surface and complementary data processing, then feeding it to the digital control unit of the CSN. Based on this, a certain robot movement is performed, such that material is adaptively applied on the constructed object. The core component of the GFS is a set of four industrial depth cameras Helios Lucid mounted on the end-effector. They are placed in specially designed IP67 certified enclosures that protect the optical system from any flying concrete particles that could potentially damage the cameras. The relevant specifications of the depth camera system are given in

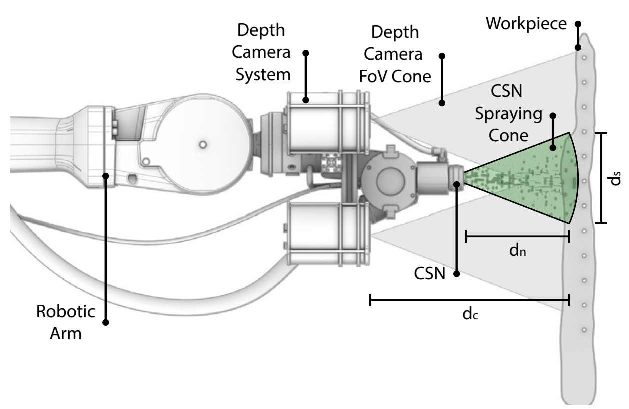

Fig. 9 Robotic setup with marked relevant components. *dc* is the scanning distance, *dn* is the distance of the digitally controlled concrete spray nozzle (CSN) to the object, and *ds* the spraying cone diameter

Table 4. The components used in the robotic spraying process are shown in Fig. 9. The central element is the robotic component, i.e. an industrial robotic arm ABB IRB4600,

Table 4 Selected specifications of the depth camera system (Lucid Vision Labs 2019); accuracy defined as measurement difference to the judgement of a skilled professional, and precision as measurement repeatability. FoV stands for field of view

FoV	59° × 45°
FoV Area	0.1 m^2 @0.4 m
Resolution	640 × 480 pixels; 0.3 MP
Range	0.3 m to 6.0 m
Accuracy	±5 mm for up to 1.5 m
Precision	0.5 mm (<0.5 m)
Acquisition rate	30 fps
Enclosure glass	AR BK7 (suitable for 850 nm λ)

which has a specified pose repeatability of 0.06 mm and pose accuracy of about 0.5 mm (ABB Robotics 2020), placed in an industrial concrete spraying facility.

4.2 3D Data Acquisition

An aspect to consider when scanning objects with depth cameras is the scanning distance, in this relation, the field of view (FoV), and the achievable resolution of scanning at that particular distance. The resolution capability of a point cloud decreases with the distance, as presented in the study of resolution capabilities for TLS in Chaudhry et al. (2021). Scanning at closer distances, however, sacrifices the size of the area of the captured object part. Nevertheless, this aspect was taken into consideration when designing the camera system to make use of four cameras without compromising an increase in the overall acquisition time. The scanning distance is fixed at approx. 40 cm to the object, thus covering 1000 cm^2 (Frangez et al. 2021a). With such configuration, we are able to distinguish surface topographic features that are larger than approximately 1 mm, a value determined based on the double the pixel area covering the object at scanning distance and as inspired by the Nyquist-Shannon sampling theorem. E.g. this value increases approximately linearly to 3 mm for a distance of 100 cm. The object is fully scanned from multiple views using various robot poses. The direct georeferencing of each individual dataset is achieved using the robot pose information as well as the hand–eye calibration parameters. These parameters give information about the translation and rotation between the tool center point (TCP) and the coordinate origin of each respective depth camera (see Frangez et al. 2021a for more information). All the acquired datasets are directly georeferenced to the robot base coordinate system, which is selected as the common coordinate system.

The data acquisition is done in a consecutive manner, i.e. once each of the production steps is completed, the surface is acquired at robot poses, which are calculated adaptively based on the object's geometry to cover the whole structure (see Frangez et al. 2021a for more information on the robot trajectory estimation). To reduce the noise in the datasets the acquisition is done in a stop-and-go way, i.e. the robotic arm is stable during the acquisition of 30 frames (i.e. 1 s in total per acquisition), before it continues to the next pose. The 30 frames are used for averaging to obtain a single dataset, i.e. a depth and an intensity image per acquisition. The first provides 3D geometric information of the object and the latter capture the power of the returned signal incorporating the radiometric properties of that same scene. An independent test is conducted to assess the impact of the camera enclosure on the noise, using a plane at a fixed distance, once with and once without the enclosure. The noise increases by 5 %, which is considered negligible and is still well within the specified precision of the depth cameras used herein (Lucid Vision Labs 2019).

The data are acquired using freshly produced concrete surfaces, therefore, understanding the implications of this on the quality of the acquisition is necessary prior to using the data in the analysis. The effects of measuring wet concrete surfaces were pointed out in the recent published literature for TLS applications (Suchocki and Katzer 2018; Garrido et al. 2019). As shown, the effect of the thin water layer on top of the concrete affects the NIR laser pulse in two ways, namely it causes (i) internal reflections and absorption of pulses within the layer of water, reducing the returned pulse power, and (ii) an increase of specular reflection of the laser beam and a reduction of diffuse one. An experiment is conducted to assess the impact of concrete wetness on the measurement noise in our setup. We measure the same quasi-flat surface of concrete, once immediately after spraying, i.e. wet, and once after curing, i.e. dry, in a quasi-perpendicular configuration. We assess the measured reflectance and the deviation with respect to a plane fit with results shown in Fig. 10. The results indicate agreement with the published research, meaning the reflectance of the wet surface is lower than that of the dry surface. Furthermore, the deviations slightly increase overall, but predominantly in the center of the image, i.e. the part that was scanned with nearly orthogonal angles of sighting within the range of approximately 10°. The latter effect is likely caused by the specular reflection, which in turn very likely caused saturation of the individual pixels within the acquired depth image. Our proposed solution to dealing with this effect is rather practical, i.e. scanning the same surface from two views, such that we are able to combine the two datasets by merging them into one, and filter out parts that were observed at sighting angles smaller than 0°. In this way, we are able to largely mitigate the effect of wet concrete onto the depth camera measurements and therefore obtain quality data for the subsequent analysis.

Further aspects to consider when performing depth camera measurements in industrial environments are predominantly related to the external temperature and its

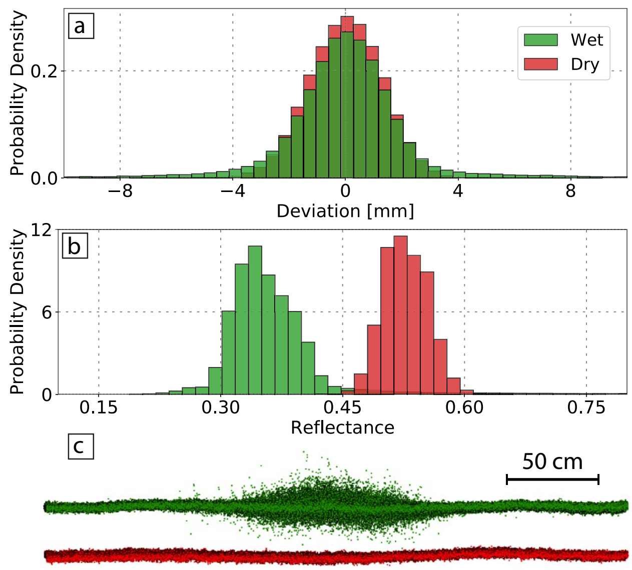

Fig. 10 Histograms showing (**a**) deviations with respect to a plane fit and (**b**) reflectances (see Sect. 3 for explanation of how it is obtained using intensity information), of a the wet and dry surface. Cross-sections (**c**) of two point clouds, one acquired on the same wet (colored green) and dry (colored red) surface

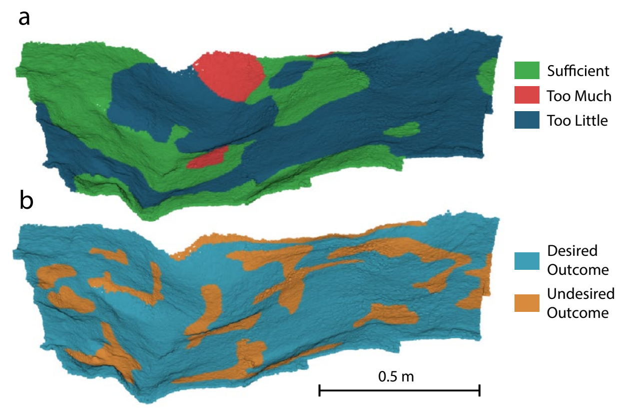

Fig. 11 Example result showing a 3D point cloud with the projected thickness map (**a**) and surface quality map (**b**). This shown dataset consists of eight individually acquired depth camera sets that are directly georeferenced to the robot coordinate system based on robot pose information

fluctuations. The outcomes of the research in Frangez et al. (2022) indicate that when external temperature fluctuations are within 2 °C, all the depth measurement errors are expected to stay within 1 mm. The temperature effect on the measurements is predominantly related to the induced mechanical and electrical variations in the camera sensor. Based on this findings, special care is taken within the production facility to assure such measurement conditions. If this is not possible, it is advised to perform a proper temperature compensation on the dataset by an independently determined correction function.

5 Results

The output of the GFS is a set of thickness and surface quality maps, which are projected onto a 3D point cloud (see examples in Fig. 11). These 3D point clouds with the two additional sets of information are passed on to the digitally controlled CSN component, which executes the needed fabrication actions. Because the 3D datasets and their corresponding maps are given in the robot base coordinate system (as enabled by the carried out hand-eye calibration procedure briefly explained in Sect. 4.2), the estimation of robot poses necessary for the action's execution can be directly estimated based on the dataset.

The prediction accuracy results are for the three datasets, i.e. manual flat, robotic flat, and robotic irregular samples, given in Table 5, for each of the three classifiers. These accuracies correspond to raw image labels without any label filtering applied. In case filtering is carried out (e.g. using a bilateral filter kernel), slight improvement in accuracy can be expected. As was the case for the manual flat samples, the surface type labels are provided by a skilled professional, such that we are able to evaluate the accuracy of our prediction. As observed, the accuracies slightly decrease when the labels are predicted on datasets acquired on wet flat surfaces that exhibited more noise than those acquired on dry samples. Further on, when applying the classifiers to the datasets acquired on irregular surfaces, the prediction decreases further, on average reaching approximately 75 % or better. The main reason for the accuracy decrease is primarily related to the highly irregular mould shape, i.e. sharp peaks and deep valleys, of the base that was used for spraying, which influenced the application of concrete. Nevertheless, the results are sufficiently good for the presented implementation of the automatic inspection and show high potential for the adoption in the production workflow.

By incorporating the presented GFS within the automatic production workflow of robotic spraying, we are able to carry out the inspection process with the sensor system capturing relevant information, and thus heavily reduce the dependence on the human in the loop. In terms of thickness

Table 5 Prediction accuracies on data acquired on manually produced samples (validation dataset) and robotically produced samples on flat and irregular surfaces for each of the three classifiers

Classifier step	Manual samples	Rob. flat samples	Rob. Irr. samples
I. Spr. GFRC	.96	.92	.86
II. Com. GFRC	.82	.80	.75
IV. Trow. Concr.	.89	.86	.80

inspection, we are able to identify areas where a corrective action is needed and determine quantitatively how much material has to be either deposited or scraped, in comparison to the very subjective nature of the human judgement. This allows for more accurate and targeted material deposition, thus increasing the sustainability of the process, without introducing redundant material deposition on the work-piece or waste. Furthermore, taking into account how labour intensive the inspection task is, it cannot be performed indefinitely by the human, while this can be the case for a robot. This in turn can have direct influence on the fabrication time, thus increasing the efficiency of the process and decreasing the amount of manual labour and thus the total production cost, especially in the long run when many pieces are produced.

The acquisition of a single image takes approximately 5 s, including the robot movement to the desired pose. The data processing of one image takes about 15 s for thickness evaluation and 20 s for surface quality using a Python script and a computer with 32 GB RAM, Intel Core i7-8086K CPU and 4.01 GHz. This performance is currently sufficient for the spraying process, since the time windows between the production steps and the time until the sprayed concrete starts to set, are on the level of 30 min. Performance optimization is possible if needed, e.g., through code optimization, parallelization and hardware upgrade.

6 Conclusion

In the paper, we presented an investigation on how 3D and intensity information provided by depth cameras can be used to automatically assess the surface quality as well as thickness of sprayed concrete elements. In particular, we proposed a pipeline incorporating local feature analysis and supervised classifiers, that incorporates different surface features and surface types, with the latter defined by the subjective judgement of a skilled professional. Further on, we have designed a geometric feedback system (GFS) that supports the robotic spraying process by outputting thickness and surface quality maps projected onto 3D point clouds.

In terms of surface quality, we were able to reach a prediction accuracy of about 85 %, using data from dry, manual flat surface samples. When the developed classification scheme was applied to robotically produced surfaces, the accuracy decreased to about 75 % or better, using data from wet, robotic irregular samples. We consider these results nevertheless as very good, since they indicate a great potential for further research in automatizing of surface inspection tasks.

The work presented in this paper is a contribution towards automatic inspection of concrete-sprayed work pieces. We presented a way of replacing subjective, slow, and costly human assessment by a fully sensor-based and automatically operating system. When implemented, it can increase the efficiency of the process, reduce costs, as well as reduce material waste of the process overall. The material depositions and surface treatment actions are targeted and based on quantitatively determined values which are reproducible and reliable, in contrast to subjective judgment of a skilled professional.

We are convinced that depth cameras are well-suited base sensors for such applications. Nevertheless consideration on incorporating other sensors, such as, e.g. RGB cameras or structured light scanners should be given in future analysis. The development presented in this paper was embedded in an interdisciplinary research project, and certain choices were dominated by specific needs and limitations of other project parts. Nevertheless, the proposed approach and solution can be applied to other construction processes as well with some adaptations which are important future work anyway. These are particularly, (i) the development of surface quality benchmark data, and (ii) strategies to react to unforeseen errors during fabrication. In this respect, one should aim to reduce the uncertainty resulting from human assessment by training the classifiers using the assessment of various professionals rather than only one. Moreover, the future research should allow for a design of desired surface types and aim at classifying relative to that design rather than classifying based on a given set of positive and negative samples. Additionally, to improve the prediction accuracy, further exploration on the relevant features that would even more deeply exploit the surface geometries of samples, should be carried out.

Funding Open Access funding provided by Swiss Federal Institute of Technology Zurich. This research was supported by the NCCR Digital Fabrication, funded by the Swiss National Science Foundation (SNSF) (#51NF40-182887) and by Innosuisse (#35019.1 IP-ENG).

Declarations

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

References

Aghdasi F, Ward RK (1996) Reduction of boundary artifacts in image restoration. IEEE Trans Image Process 5(4):611–618

Bard J, Bidgoli A, Chi WW (2018) Image classification for robotic plastering with convolutional neural network. Robotic fabrication in architecture, art and design. Springer, Berlin, pp 3–15

Buitinck L, Louppe G, Blondel M, Pedregosa F, Mueller A, Grisel O, Niculae V, Prettenhofer P, Gramfort A, Grobler J, Layton R, VanderPlas J, Joly A, Holt B, Varoquaux G (2013) API design for machine learning software: experiences from the scikit-learn project. In: ECML PKDD workshop: languages for data mining and machine learning, pp 108–122

Buswell R, Kinnell P, Xu J, Hack N, Kloft H, Maboudi M, Gerke M, Massin P, Grasser G, Wolfs R et al (2020) Inspection methods for 3d concrete printing. RILEM international conference on concrete and digital fabrication. Springer, Berlin, pp 790–803

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

DIN EN ISO 4287:2010-07 (2010) Geometrical Product Specifications (GPS) - Surface texture: Profile method—Terms, definitions and surface texture parameters. Berlin, De, Standard, [Online]. Available: https://doi.org/10.31030/1699310

Ercan Jenny S, Lloret-Fritschi E, Jenny D, Sounigo E, Tsai P-H, Gramazio F, Kohler M (2021) Robotic plaster spraying: crafting surfaces with adaptive thin-layer printing. 3D Print Addit Manuf 2:2

Eversmann P, Gramazio F, Kohler M (2017) Robotic prefabrication of timber structures: towards automated large-scale spatial assembly. Constr Robot 1(1):49–60

Fooladgar F, Kasaei S (2020) A survey on indoor rgb-d semantic segmentation: from hand-crafted features to deep convolutional neural networks. Multimed Tools Appl 79(7):4499–4524

Frangez V, Lloret-Fritschi E, Taha N, Gramazio F, Kohler M, Wieser A (2021) Depth-camera-based rebar detection and digital reconstruction for robotic concrete spraying. Constr Robot 2:1–12

Frangez V, Salido-Monzú D, Wieser A (2021) Surface finish classification using depth camera data. Autom Constr 129:103799

Frangez V, Salido-Monzú D, Wieser A (2022) Assessment and improvement of distance measurement accuracy for time-of-flight cameras. IEEE Trans Instrum Meas 71:1–11

Garrido I, Lagüela S, Sfarra S, Madruga F, Arias P (2019) Automatic detection of moistures in different construction materials from thermographic images. J Therm Anal Calorim 138(2):1649–1668

Hadavandsiri Z, Lichti DD, Jahraus A, Jarron D (2019) Concrete preliminary damage inspection by classification of terrestrial laser scanner point clouds through systematic threshold definition. ISPRS Int J Geo Inf 8(12):585

Hassaballah M, Abdelmgeid AA, Alshazly HA (2016) Image features detection, description and matching. Image feature detectors and descriptors. Springer, Berlin, pp 11–45

Jenny SE, Lloret-Fritschi E, Gramazio F, Kohler M (2020) Crafting plaster through continuous mobile robotic fabrication on-site. Constr Robot 4(3):261–271

Kazemian A, Yuan X, Davtalab O, Khoshnevis B (2019) Computer vision for real-time extrusion quality monitoring and control in robotic construction. Autom Constr 101:92–98

Liu Y, Li Z, Liu H, Kan Z (2020) Skill transfer learning for autonomous robots and human-robot cooperation: a survey. Robot Auton Syst 128:103515

Maboudi M, Gerke M, Hack N, Brohmann L, Schwerdtner P, Placzek G (2020) Current surveying methods for the integration of additive manufacturing in the construction process. Int Arch Photogram Remote Sens Spat Inf Sci 43:763–768

Mineo C, Vasilev M, Cowan B, MacLeod C, Pierce SG, Wong C, Yang E, Fuentes R, Cross E (2020) Enabling robotic adaptive behaviour capabilities for new industry 4.0 automated quality inspection paradigms. Insight-Non-Destr Test Cond Monitor 62(6):338–344

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830

Product specification IRB 4600, ABB Robotics (2020)

Rodríguez-Gonzálvez P, Guidi G (2019) Rgb-d sensors data quality assessment and improvement for advanced applications. RGB-D image analysis processing. Springer, Berlin, pp 67–86

Schindler K (2012) An overview and comparison of smooth labeling methods for land-cover classification. IEEE Trans Geosci Remote Sens 50(11):4534–4545

Speiser JL, Miller ME, Tooze J, Ip E (2019) A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst Appl 134:93–101

Suchocki C, Katzer J (2018) Terrestrial laser scanning harnessed for moisture detection in building materials-problems and limitations. Autom Constr 94:127–134

Sutjipto S, Tish D, Paul G, Vidal-Calleja T, Schork T (2018) Towards visual feedback loops for robot-controlled additive manufacturing. Robotic fabrication in architecture, art and design. Springer, Berlin, pp 85–97

Taha N, Walzer AN, Ruangjun J, Bürgin T, Dörfler K, Lloret-Fritschi 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

Time-of-flight forged ahead: design tips to boost 3D performance and cut integration time and cost (v.1.3), Lucid Vision Labs (2019)

Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Sixth international conference on computer vision (IEEE Cat. No. 98CH36271). IEEE, pp 839–846

Wangler T, Roussel N, Bos FP, Salet TA, Flatt RJ (2019) Digital concrete: a review. Cem Concr Res 123:105780

Weinmann M, Jutzi B, Mallet C (2014) Semantic 3d scene interpretation: a framework combining optimal neighborhood size selection with relevant features. ISPRS Ann Photogram Remote Sens Sp Inf Sci 2(3):181

Wolfs R (2019) Experimental characterization and numerical modelling of 3d printed concrete: controlling structural behaviour in the fresh and hardened state. Ph.D. dissertation, Built Environment, proefschrift

Wolfs RJ, Bos FP, van Strien EC, Salet TA (2018) A real-time height measurement and feedback system for 3d concrete printing. High tech concrete: where technology and engineering meet. Springer, Berlin, pp 2474–2483

Yao X, Han J, Cheng G, Qian X, Guo L (2016) Semantic annotation of high-resolution satellite images via weakly supervised learning. IEEE Trans Geosci Remote Sens 54(6):3660–3671

Zaczek-Peplinska J, Osińska-Skotak K (2018) Concrete surface evaluation based on the reflected tls laser beam's intensity image classification. Stud Geotech Mech 40:1