Depth-Camera-Based In-line Evaluation of Surface Geometry and Material Classification For Robotic Spraying

V. Frangez, D. Salido-Monzú, and A. Wieser

Institute of Geodesy and Photogrammetry (IGP), ETH Zürich, Switzerland

valens.frangez@geod.baug.ethz.ch, david.salido@geod.baug.ethz.ch, andreas.wieser@geod.baug.ethz.ch

Abstract -

This paper presents a feasibility study of surface geometry (SG) evaluation and material classification (MC) for robotic spraying. We propose two complementary approaches using point clouds and intensity data provided by a state-of-the-art industrial time-of-flight (ToF) depth camera. The SG evaluation is based on geometric feature computation within local neighbourhoods, which are then used within a supervised classification. The results of this approach are SG classes according to the level of geometric variability of the surface, displayed as SG maps. For MC, active reflectance estimation is investigated and exploited to derive features related to the reflectance and diffusive properties of each material for classification. The result of both approaches can be prospectively used as feedback in digital fabrication for in-line adaptation of the process to improve control of relevant geometrical and material properties.

Keywords -

Surface Geometry; Material Classification; Digital Fabrication; Depth Camera; Machine Learning

1 Introduction

Assuring the required geometric quality of digitally fabricated structures is of high importance not only after but also during the construction process. Recent rapid growth of interest in robotically assisted construction [1] has also boosted the use of sensing technologies to acquire process-relevant information. These works showcased the potential of various sensors in digital fabrication processes and demonstrated ways in which accurately measured 3D information and object parameters extracted from such information can be used for in-line process improvement via feedback control [2, 3]. In order to extract the relevant information for construction processes, appropriate processing algorithms and interpretation of the 3D data need to be employed.

In this contribution we focus on the concept development and feasibility study of algorithms for surface geometry (SG) evaluation and material classification (MC) for robotic spraying applications. The goal is to address the possibility of replacing the manual and subjective SG evaluation and MC processes, currently relying on experts, with novel sensing technology and data analysis. A state-of-the-art industrial depth camera was chosen for data acquisition within this investigation. This sensor technology offers a good trade-off between in-line process acquisition capacity, sufficiently high accuracy and resolution, and areal coverage on the object of interest. It provides not only geometric information of the acquired scene in the form of 3D point clouds (PCs), but also the underlying raw data related to the measured distances and intensities. These data types have potential for extracting various relevant parameters of the observed object by employing a single sensor solution.

SG of structures produced with spraying is important not only for visual and design considerations, but also from a structural viewpoint. The acquired PC data follows the SG of the captured object and implicitly contains its geometric properties. To exploit these geometric properties, i.e. features, it is necessary to carry out a 3D analysis within local neighbourhoods. To achieve this, a processing scheme of four main steps is proposed in this paper: (i) separation of the acquired SG to spatial frequency components, (ii) neighbourhood selection, (iii) feature computation, and (iv) supervised classification. The outcome of this approach are geometry classes according to the level of geometric variability of the surface for different spatial frequencies, displayed as SG maps.

Obtaining adequate information of the early hydration state or local coverage and volume of each sprayed material (e.g. concrete, plaster finish, or paint) requires classifying the materials within the observed scene. This is critical for closed-loop control of the spraying process, such that complete coverage of the surface per material can be ensured. In this context we explore MC exploiting the observations provided by a time-of-flight (ToF) depth camera related to material-specific reflectance properties. The observed differences on repeatable features for the materials used in this initial investigation (i.e. wood, cardboard, metal, and plastic) are then used for developing a simple classifier based on a k-nearest neighbour (kNN) algorithm.

The contributions of this study are twofold. On the one hand, the SG evaluation algorithm can be used within the geometric feedback system of the fabrication in order to adapt the process accordingly. Chosen SG class ought to be considered as a goal SG, while others indicate areas that need further surface treatment, i.e. either adding additional geometry variation or applying more material to smooth out the dominant SG features. On the other hand, the MC approach can be used within a process of early concrete hydration by distinguishing different states of the concrete and their associated reflectances to automatically regulate the time-window between consecutively sprayed layers.

To introduce this study, a brief overview of the working principle of ToF depth cameras and a basic performance assessment of the one used within this work is given in Section 2. The specific SG evaluation (Section 3) and MC (Section 4) approach is then detailed, including an overview of processing steps, used experimental setups, and datasets. The results of the two discussed concepts are then given and interpreted in Section 5, and the general conclusion and outlook of the investigation are summarized in Section 6.

2 ToF camera

2.1 Measurement Principle

Depth cameras produce 2D depth images of a scene relying on various measurement principles, such as stereo triangulation, structured light, or ToF approaches [4]. With the increase in sensor resolution, the latter have gained attention in recent years for being able to operate over larger ranges with accuracy levels more independent of the scene geometry and background illumination. The most established technique to implement ToF cameras is based on measuring the accumulated phase of an intensity modulated optical signal illuminating the scene and detected independently on each pixel. The illumination is typically produced by one or several near-infrared LEDs driven by a radio-frequency (RF) carrier. The detected signal on each pixel, associated to the illumination signal back-reflected on the corresponding patch of the scene, is demodulated at device level by a dedicated 2D array that computes the correlation of the received signal with an internal reference derived directly from the driver of the emitted modulation. Simplifying the modulation to one sinusoidal component, this correlation $c(\tau)$ can be written as

$$c(\tau) = a \cdot \cos(\tau + \phi) + b, \tag{1}$$

where τ is the correlation lag, a the optical intensity at the modulation frequency, ϕ the phase shift proportional to the propagation delay—and therefore to the range of interest—and b the illumination bias. This correlation is typically computed for four lags regularly spaced with $\frac{\pi}{2}$ rad, so that the obtained correlation samples A_i are

$$A_i = c\left(i \cdot \frac{\pi}{2}\right), \quad \text{with } i = 0, \cdots, 3. \tag{2}$$

The distance-dependent phase can be directly calculated from this correlation samples as

$$\phi = \arctan\left(\frac{A_3 - A_1}{A_0 - A_2}\right) \tag{3}$$

which, considering a sufficiently well-known value of the modulation frequency f_m and assuming a sufficiently accurate approximation for the propagation speed c of the optical signal under typical operation meteorological conditions, can be used to estimate the distance of interest for the given pixel as

$$\hat{d} = \frac{c}{4\pi f_m}(\phi - \phi_0) \tag{4}$$

where ϕ_0 accounts for the systematic delay offset in the correlation process with respect to the mechanical zero of the camera, which should be compensated by calibration. As any phase-based measurement system, these observations are ambiguous beyond one modulation wavelength. This ambiguity is practically solved by performing quasi-simultaneous measurements at several f_m and combining them to resolve a larger unambiguous range. Once computed for each pixel of the depth image, the estimated distances can be used to generate a 3D PC by means of perspective projection [5].

In addition to the accumulated phase, the optical intensity a can also be computed from the correlation samples as

$$a = \frac{1}{2}\sqrt{(A_3 - A_1)^2 + (A_0 - A_2)^2}, \tag{5}$$

being typically used to augment the generated PCs with intensity values. These intensities are proportional to the amplitude of the detected modulated signal, hence independent of the ambient illumination. Given a sufficiently stable illumination and detection efficiency, the measured intensities can be processed considering the spatial distribution of the illumination signal and radiometric attenuation with distance and viewing angle, what enables deriving observations proportional to the material reflectance independently on its position on the scene. By including a prior calibration of scale using a known reflectance standard, these compensated intensities can be in turn extended to absolute reflectance estimations for each point on the scene. Note that such estimations would correspond to the reflectance of the illuminated patch of surface associated to each pixel at the specific angle of incidence (AOI) defined by its relative orientation with respect to the camera.

2.2 Performance Evaluation

An evaluation of the performance using a state-of-the-art ToF camera Helios Lucid [6], has been carried out. The camera operates based on the measurement principle described in Section 2.1, and its most relevant specifications are given in 2.2. This camera provides both raw correlation samples as directly computed on each pixel according to (2), and on-device pre-processed PC data after averaging multiple frames, correcting of systematic errors, and spatial data smoothing to reduce noise. The ToF sensor used in the camera allows users to access raw data for four modulation frequencies f_M (25 MHz, 50 MHz, 75 MHz, and 100 MHz), however for the processed PC dataset, only 75 MHz and 100 MHz are employed as indicated in Table 2.2.

Table 1. Selected relevant properties of Helios Lucid depth camera, see [6] for more.

Property	Value
Resolution	640 pix x 480 pix
FoV	59° x 45°
Mode	1.5 m (75 MHz and 100 MHz) 6 m (100 MHz)
Precision	<1.6 mm @ 1 m
Accuracy	± 5 mm (1.5 m mode) ± 10 mm (6 m mode)
Illumination	850 nm (VCSEL laser diodes)

Practical limitations on the achievable accuracy of ToF depth cameras are mainly related to sources of variability for the systematic deviations, such as changes of the operating conditions, namely environment temperature, or multipath effects on the scene [5]. Performance assessment procedures of commercial depth cameras primarily addressing those limitations are presented in [7, 4]. Some of these error sources can be mitigated to a large degree by calibration and others further reduced by a careful design of the acquisition conditions and geometry when allowed by the application. Additionally, depth precision is dominated by the signal quality at the sensor level, proportional to the detected optical power and therefore defined mostly by the distance and relative orientation between the sensor and the scene patch, and the reflectance properties of the latter. Precision can only be improved under certain acquisition conditions at the expense of dynamic performance for the given application, by averaging over several measurements of an unchanged scene.

Before using our camera for the data acquisition of our proposed application, We have performed a basic performance assessment to gain a better understanding of the expected data quality and practical limitations. This assessment has specifically addressed stability and noise figure, warm-up effect, and intensity-related depth errors.

The stability and noise characteristics of the raw data were evaluated by computing the Allan deviation of the phases calculated according to (3). This computation was based on a time-series of 8000 samples acquired over two hours on a static scene composed of a flat white plane at a distance of about 1 m using the 1.5 m operating mode. Phase stability was analysed for each of the f_M, using values for a single pixel and is shown in Figure 1. The results indicate a high long term stability on the evaluated time scale without significant low-frequency drifts. A linear fit based on the Allan deviation yields a slope very close to -1, indicating that the noise background is mostly uncorrelated and flat over the evaluated bandwidth—the expected improve by averaging over several acquisitions is therefore approximately proportional to the square root of the number of acquisitions. Computing the distance corresponding to each phase according to 4, the measured precision (2σ) for a single acquisition on each frequency on the evaluated conditions are 5.5 mm (25 MHz), 3.3 mm (50 MHz), 2.9 mm (75 MHz) and 2.3 mm (100 MHz). Precision is expected to increase proportionally to the modulation frequency due to the reduction of distance sensitivity to phase errors. This increase is, however, partly limited by a slight reduction of detected power on faster modulations more attenuated by the sensor bandwidth.

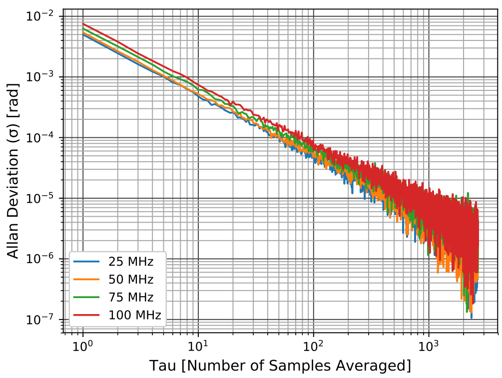

Figure 1. Allan deviation of phase observations for each f_M, using a sample of 8000 measurements for a single pixel.

The internal warm-up of the camera and its impact on accuracy due to temperature-dependent scale errors was evaluated to determine empirically the required warm-up time until reaching optimal performance. Static measurements under similar conditions as described above for precision were carried out acquiring PCs on a flat white plane. This process was initiated immediately after the device was turned on and carried out over three hours with measurements triggered every 20 s. Each PC was used to fit a plane, based on which the maximum and average deviations between the points and the fitted plane were computed. To reduce the impact of outliers in the analysis, the maximum deviations were computed as the median of the largest 20 deviations. Initial maximum and average deviations were 14 mm and 6 mm, respectively. The results indicate the camera performance stabilizes at about 45 minutes after the camera was turned on, i.e. when the internal camera temperature reaches 36°C. The maximum and average deviations then drop and remain constant at 8 mm and 2 mm, respectively.

The evaluation of intensity-related depth errors due to the spatial inhomogeneity of the illumination pattern were assessed by acquiring PCs using the 6 m operating mode on a white flat surface at distances between 30 cm and 120 cm on 10 cm steps. Deviations to a plane fitted to each PC were computed for each distance. The deviations were grouped in three categories: deviations up to 2 mm, deviations from 2 mm to 4 mm, and deviations larger than 4 mm. These values were chosen to approximate the empirically computed precision discussed in the previous section and its double. The results can be seen in the Figure 2 (left). The figure also shows the area covered by the camera at selected distances. An oscillation patter can be seen within the deviation groups. This pattern agrees with oscillations observed in the distances computed using the raw phase data at 100 MHz (not shown), which suggest a residual uncalibrated cyclic error. Based on the obtained results, it is advised to acquire data at either 40 cm or 80 cm, where the deviations are minimum. Figure 2 (right) shows PCs colored according to the previously defined deviation categories. The most and least optimal PCs were chosen for this representation, being 40 cm and 60 cm respectively. The observed errors could be further mitigated by applying an online compensation taking into account the measured intensity and distance. As a simple approach for the data acquisition during the presented investigation, we have designed our experiment setup geometry and correspondingly cropped the acquired data to exclude the areas with highest deviations.

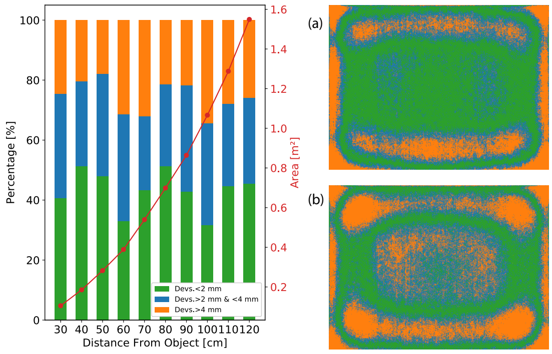

Figure 2. Deviations grouped in three categories, displayed in proportions. Red line represents the area covered by the camera as a function of distance (left). PCs acquired at 40 cm (a) and 60 cm (b) colored according to deviation category (right).

3 Surface Geometry Evaluation

SG captured within the PC can be exploited by local geometric feature analysis. A proposed scheme of assigning points a class of geometry according to the power of change is depicted in Figure 3. The input point cloud is first pre-processed. Then the data is partitioned to three spatial frequency components, and based on a selected neighbourhood, features are computed for each point within the dataset. Next a supervised classification is carried out, outputting the classes for each point which are then spatially filtered. The outcome of this six step processing are SG classes displayed as SG maps.

In general, the presented approach relies on the input data used for model training in the classification step. The observed geometry in the investigated samples is not referenced to a certain absolute definition, but grouped together according to distinctive SG variations.

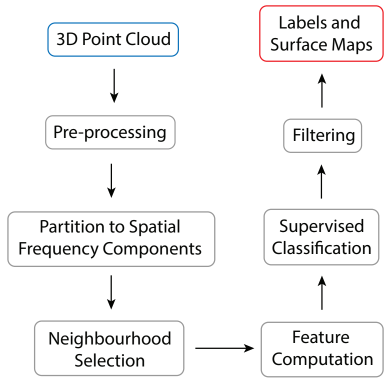

Figure 3. Proposed processing pipeline for SG evaluation.

For the purpose of concept development and evaluation, concrete samples of various geometries were produced using mould casting. Six out of seven sample geometries are shown in Figure 4. These types of geometry were chosen such that they resemble the sprayed surface, since the actual sprayed surface samples were not yet available for this investigation. The chosen dataset is however not completely representative of the sprayed surface and as mentioned, changing the input dataset would lead to different class definitions. Variations in height components (based on Reckli GmBh specifications) are from 11 mm up to 60 mm, therefore covering a wide range of geometric diversity.

The first step after acquiring the data is to pre-process it, which includes two operations: data detrending and cropping. The first operation removes the global linear trend from the whole PC using principal component analysis and reorients it according to the first and second greatest variance of the coordinates. The measured surface might be oriented and at certain angle with respect to the camera, leading to wrong assumptions about a certain geometry. The second operation is to crop the PC and to keep only the central part, which is less affected by intensity related errors compared to the edge parts and therefore displaying higher accuracy (see Section 2.2).

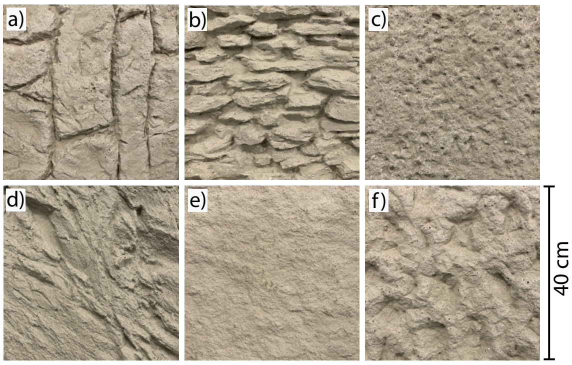

Figure 4. Concrete samples used within the SG evaluation investigation.

3.1 Partition to Spatial Frequency Components

A systematic approach for SG evaluation is adopted based on an existing method for inspection in surface metrology [8], which separates the surface into form, waviness, and roughness components. Typical values for these three components are within a range from microns to a few mm. We have adapted this concept within our SG evaluation and made it representative of the level of geometric variations present in our samples that is both relevant for digital fabrication and observable with the proposed technology, covering a range from a few mm to several cm. We established relative terms, which cover the complete geometric range of the used sample data, namely low- (LSFG), medium- (MSFG), and high-spatial frequency geometry (HSFG).

Two approaches of partitioning the SG to components were employed, making use of either discreet Fourier (FT) or discreet wavelet transform (WT). Since the PC structure is a grid, the transformation can be directly applied to columns and rows of the dataset without any need for interpolation of points. A spatial profile is treated as a time series signal where, instead of time, X- or Y-coordinates are used for rows and columns, respectively. In the first approach, the frequency threshold values between the components of the spectrum were determined empirically based on the observed spectrum in the frequency domain after using the FT. Then Butterworth low-, band-, and high-pass filters were used to partition the signal into three components. An empirically optimized order of the filter was used, which led to more optimal results. In the second approach, using the WT, a selection of a wavelet family and a decomposition level had to be made. Coiflet wavelets and level 5 were empirically selected, since they showed the most optimal results, after trying out different parameter options. The first parameter defines the signal that will be used to convolute the spatial profile, while the second sets the number of levels (i.e. components) that the signal gets partitioned to. In our case LSFG corresponds to level 1, MSFG to levels 2 and 3, and HSFG to levels 4 and 5.

The results of the two approaches applied to a randomly selected spatial profile from surface sample c) are shown in Figure 5. The tree SG components, the original SG, and the reassembled SG are shown. It can be noticed that the results based on the FT approach cannot be reassembled to the original SG profile, while this is successfully achieved for the WT. This confirms that Wavelets perform better on natural signals, i.e. non-stationary signals, as demonstrated in [9]. The FT approach using Butterworth filters does not, on the other hand, perform well on sharp changes in the signal and it tends to smooth over those features.

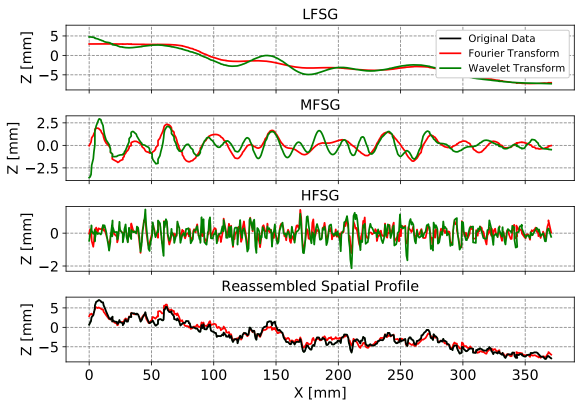

Figure 5. Profiles for each of the three surface components for surface sample c).

The whole PC can be partitioned to components, once this approach is applied to the complete dataset, i.e. all rows and columns. Both approaches are used in this processing, namely WT for partitioning to SG components (see example in Figure 7), and FT for generation of spectrograms. Spectrograms are a visual representation of the spectrum of frequencies of the signal in the time domain, however here shown for spatial domain. A visualization for surface sample b) is shown in Figure 6. The visible information in the spectrograms differs for the dataset obtained from geometries as shown in Figure 4. It can serve to provide additional information of the spatial frequency content of the surface to the user.

The intensity image is obtained by using the same WT approach partitioned to a high- (HFI) and low-frequency intensity (LFI). The LFI component is dominated by the contribution of the camera illumination pattern and is discarded from further analysis. The HFI component, on the other hand, is influenced by small scale features in the observed surface, and can be therefore used to extract small scale features.

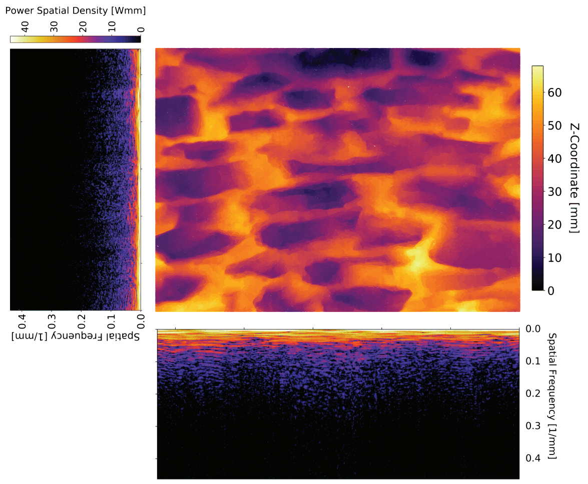

Figure 6. Spectrograms generated for surface sample b) in direction of columns and rows.

Only WT results were used in the end for this investigation, since the results were more optimal compared to the FT. The FT and its results were only used for generation of the SG spectrograms.

3.2 Geometric and Intensity Feature Computation

Each surface component dataset is used for extraction of local geometric features, with features being different for each of the three components. The selected features were handpicked in order to capture the information within the SG components as good as possible. The selection is shown in Figure 7. More information on geometric features that were taken as a starting point in this investigation and on their properties can be found in [10].

The computation of features is carried out for each point within the dataset, using its neighbourhood. Within the implemented algorithm, points are included within a certain neighbourhood based on either the list of indices of the kNN of the anchor point or a query of all points with distances to the anchor point smaller than a given radius. The results shown in this paper were produced based on a chosen radius, being 7 cm, 3 cm, and 1 cm, for LSFG, MSFG, and HSFG, respectively. This makes the classification results independent on the distance from which the acquisition was done.

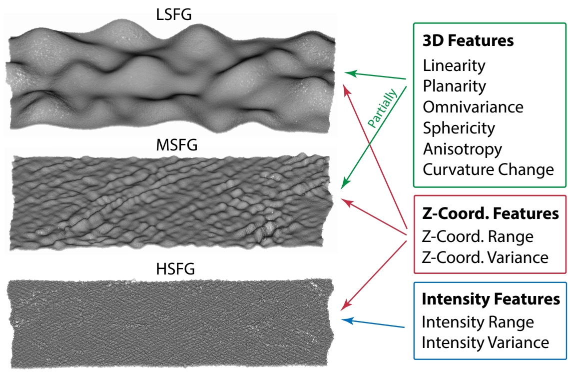

Figure 7. Three SG components (left). A list of features used for computation (right).

The intensity features play a relevant role when analyzing the HSFG component. The goal is to exploit small scale (i.e. sub-mm) features that cannot be captured by the geometry of the PC, due to the accuracy limitation of the used technology. However, the input intensity images of highly reflecting and absorbing local areas could cause intensity variations, in this case the intensity variations would not be related to the geometric but rather to radiometric features.

3.3 Supervised Classification

Once the features for all points are computed, they are normalized per feature to a range of [0, 1], such that none of the feature values dominate on the further processing. Those values are then used within a process of clustering to form groups of points that share similar geometric properties. For this purpose, an agglomerative hierarchical clustering approach was chosen. The algorithm requires a user specified number of clusters. This can be chosen based on a dendrogram, a tree diagram that displays an arrangement of clusters and the similarity between them. The number of clusters was selected as 3 for all components. Only a random subset of about 10% of points with their features was used for model training. It is assumed that the sub-set is sufficiently representative for the whole dataset, which was validated by the result repeatability after selecting different random subsets. The rest of the points is used later on for the evaluation of the model.

Next, the segmentation approach is extended to supervised classification using kNN. The algorithm learns the relationship between the feature value list of each point and assigned cluster indices. The number of clusters from the clustering process and the classification process is therefore the same. This classifier can then be applied to another dataset, to predict the closest class label using kNN.

To derive labeling of points with higher spatial regularity, a median filter is applied to the initial class label image. This results into smooth labels and surface maps. Additionally, it is possible to omit patches of areas smaller than the chosen threshold to decrease granularity, however this is not implemented in the proposed approach.

Furthermore, an assessment of each of the features was carried out to understand which have the highest impact on the label outcome. For this the analysis of variance (ANOVA) statistical test was employed, usually used within similar classification tasks. Based on the features and assigned labels, ANOVA assesses how much a class label is dependent on each particular feature and assigns each feature a score of relevance. It then ranks the features based on the calculated score. The most important features were intensity range and Z-coordinate range for HSFG and MSFG, respectively. For LSFG the most important features were change of curvature, omnivariance, and sphericity. The results of the labeling are shown in Section 5.1.

The process of feature computation and classification is computationally intensive, since a set of feature values has to be computed for each point. This makes it unsuitable for a quasi-real-time application. However, this can be overcome by computing features for a subset of points, e.g. every 10^{th} point, classifying them, and then using interpolation of classes on the rest of the points to produce fully classified maps.

4 Material Classification

We have investigated the MC potential of ToF cameras in the context of our targeted applications by developing a MC strategy exploiting active reflectance estimations derived from the raw correlation data. The goal is to propose a processing pipeline to derive information of the reflectance and diffusiveness of the materials in the scene and use it for in-line classification, and to provide a simple demonstration of the feasibility of such approach.

4.1 Experiments and Datasets

An experimental setup for material-related data acquisition was designed making use of the horizontal comparator bench on our calibration laboratory. The ToF camera was mounted on a steel pole in front of the bench and the material samples were placed on a motorized trolley that moves automatically between selected positions measured with a linear Doppler interferometer. This allowed for accurate control of the relative distance changes between samples and camera, which were extended to absolute distances by independently measuring the distance between the camera and the sample at the starting position of the trolley using a laser tracker.

Two experiments were carried out on the described setup using four material samples, namely wood, plastic, metal, and cardboard, chosen arbitrarily as a simple subset of common materials. In the first experiment, each of the four material samples was placed on the trolley and displaced from 0.57 m to 1.6 m on steps of 3 cm. The scene on each position was acquired 10 times sequentially to enable an approximate quantification of measurement dispersion. The complete procedure was repeated for three orientations of the samples (5°, 25°, and 35°) by controlling the orientation with a highly repeatable rotation stage. The values for the distance range and sample orientations were selected to cover the expected working regions of the camera for the envisioned applications, providing sufficient resolution while keeping the effort for the experiments low. In the second experiment, all four material samples were placed on the trolley. The scene containing all materials was then captured for distances between 0.57 m and 1.07 m on steps of 0.1 m, repeating the process for the three aforementioned orientations.

4.2 Active reflectance estimation

Our proposal for reflectance-based MC relies on distinguishing materials based on their reflectance pattern. This approach is based on the estimation of a simple model of the material absolute reflectance as a function of AOI, obtained from calibrated intensity observations of the same surface patch from different points of view. The features extracted from such estimation can potentially provide differentiation for a large number of materials as long as they do not showcase both similar absolute reflectance and diffusive behaviour.

Defining in the simplest case a two-parameter reflectance model R as a generalized Lambertian scatterer

$$R\,(\mathrm{AOI}) = R_0 \cdot \cos^n(\mathrm{AOI}), \tag{6}$$

where R_0 is the material absolute reflectance on normal incidence and n defines the patter directivity; materials can be characterized based on their features R_0 and n. These are obtained from a minimum but sufficient number of two observations of the same surface patch from different perspectives, each observation providing estimations of both the absolute reflectance and the AOI. The AOI can be estimated from the depth map by analyzing the local region of the pixel of interest. The absolute reflectance, on the other hand, can be derived from the intensities calculated from the correlation raw data by applying adequate distance and instrumental corrections. The intensity a measured at pixel px for certain distance d, considering the dominant sources of variability, can be modeled as

$$a\,(d, px) = K_{ill}\,(d, px) \cdot K_{det}\,(px) \cdot G\,(d) \cdot R\,(\mathrm{AOI}), \tag{7}$$

where K_{ill} accounts for the spatial distribution of the illumination signal, K_{det} represents the detection efficiency of the corresponding pixel, and $G(d)$ the overall radiometric attenuation of detected power with distance. These three components can be compensated online using fitted functions or lookup tables obtained by a single calibration process based on acquisitions of a known geometry (e.g. a planar target) with homogeneous reflectance properties over several distances within the range of interest. By additionally including measurements on a material with known reflectance such as a calibrated reflectance standard, the overall scale of the amplitudes can be corrected. The final compensated amplitude provides an estimation of the material reflectance $\hat{R}$ of the surface patch covered by each pixel independently of distance and viewing angle from the camera, thus enabling the generation of absolute reflectance images. Combining these images with the estimated AOI per pixel for pairs of acquisitions allows deriving and mapping the reflectance features R_0 and n for any part of the surface covered on both acquisitions. These features can be then used to classify the materials in the scene by direct comparison with a pre-collected material database.

We have used the data from the two experiments described in Section 4.1 to provide a simple validation of this approach. The correlation data from the first experiment was used to compute the optical intensities per modulation wavelength according to (5). Figure 8 shows an extract of the computed intensities as a function of the interferometric reference distance across all the evaluated materials for the fastest modulation wavelength, with solid lines representing the average of 10 acquisitions per position while shaded areas correspond to the dispersion of those acquisitions in $\pm 1\ \sigma$. These results are good representatives of the overall behaviour of the other three f_M. The analysis of the intensity values indicates unreliable intensity observations due to saturation for some materials at close range. This is specially critical on the metal sample that showcases much higher specular reflectance. This makes the received intensity very high for the 5° orientation, which saturates the sensor for any distance smaller than approximately 1.1 m as confirmed by inspection of the underlying correlation samples. Conversely, the more specular behaviour produces very low intensities for the AOIs on the 25° and 35° orientations. The data derived from saturated observations was excluded from subsequent analyses.

The intensities obtained in the first experiment have been used to generate reference data for the classification. To avoid the need of introducing a reflectance standard, the classification is carried out using relative reflectance values normalized to the average amplitude of wood at 5°. The computed relative reflectances are depicted in Figure 9 for the three evaluated AOIs, where solid and

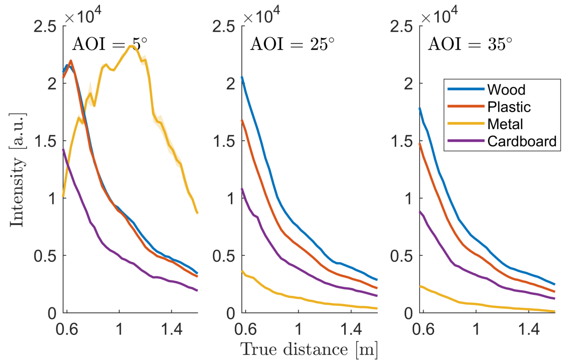

Figure 8. Measured intensity at 100 MHz as a function of reference distance for all materials and AOIs.

shaded areas represent the average and standard deviation ($\pm 1\ \sigma$) between all wavelengths. The results show that, excluding the regions when the sensor is close to saturation due to high material reflectance and short range, the computed values per material are consistent within a 10% of variability independently from distance. This indicates that features derived from these values hold potential to provide robust classification of the evaluated materials.

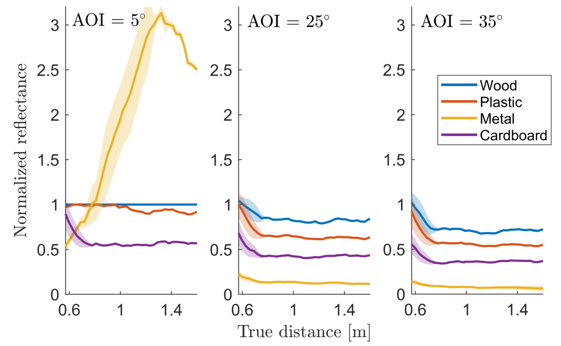

Figure 9. Normalized intensities at all f_M as a function of reference distance for all materials and AOIs.

The relative reflectances per AOI computed for each material, f_M, and distance have been used to estimate the two parameters R_0 and n of the simple reflectance model in (6), which are to be used as reference features for the classification. Figure 10 shows the resulting features for all frequencies and distances with their associated material class. As seen in the figure, the computed features provide good separability for the evaluated materials.

The data from the first experiment has been additionally used to derive the distance correction $G(d)$, and an additional independent experiment on a planar target at several distances has been used to compute the illumination distribution $K_{ill}(d, px)$. These corrections are used on a scene containing all materials from the second experiment to test

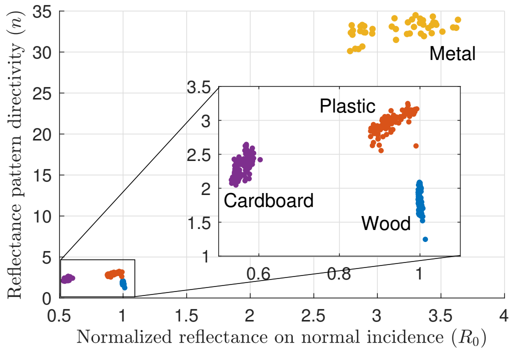

Figure 10. Normalized reflectance features extracted for four materials for all frequencies and distances.

the MC performance against the reference data. The intermediate steps for these corrections and the classification results are shown and discussed next in Section 5.2.

5 Results

5.1 Surface Geometry

The results of the SG evaluation are maps colored according to the assigned classes. Each class signifies the level of geometric variability for each of the spatial frequency components LSFG, MSFG, and HSFG (Figure 11).

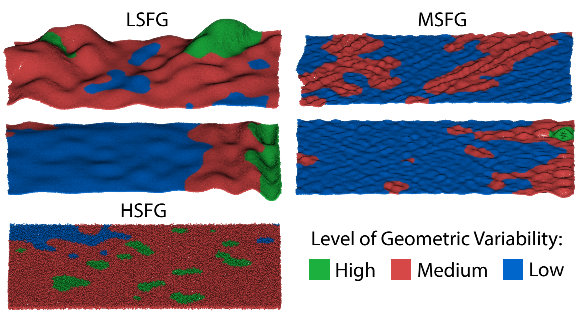

Figure 11. Resulting classified SG map for LSFG, MSFG, and HSFG components.

The results indicate that the proposed evaluation approach works well, with clearly classifying the most similar geometric parts of the datasets in the same class. The differences between classes are obvious and very distinctive. The level of in-class variation is directly related to how many classes were chosen in the classification step. For the given dataset this value has been optimized, to allow only for minor variations within the classes. The selection of the number of the classes should thus be a data driven decision. Due to the unavailability of ground truth data, a plausibility check of the obtained results was done based on subjective evaluation. This resembles standard practice, where SG is evaluated by human inspection.

The results are useful within the process of robotic spraying, where certain classes indicate areas requiring further surface treatment. If the target SG would be a class of medium geometric variability, regions showing high variability imply the need of being filled with more material around it or removed, while regions showing low variability require additional texture.

5.2 Material Classification

A scene containing all four materials seen from two different perspectives (5° and 25° with respect to the camera axis) has been used to assess the MC performance. The intensities as directly computed from the correlation samples are shown in Figure 12.a, for wood (W.), cardboard (C.), plastic (P.), and metal (M.). These intensities were compensated for the impact of distance and illumination spatial distribution with the corrections calculated as described in Section 4.2. The illumination distribution can be seen in Figure 12.c, and the resulting distance- and illumination-corrected intensities are shown in Figure 12.b. The AOI across the scene was calculated geometrically, and the need for calibrating the detection efficiency and overall scale was overcome by computing reflectances normalized to a patch of the scene known to correspond to wood on the 5° acquisition. The two reflectance features were then derived from the computed relative reflectances and compared to the reference data using a kNN classifier. The final results of the classification per pixel after applying a median filter can be seen in Figure 12.d. The results demonstrate high accuracy on the classification of pixels corresponding to metal, cardboard, and wood. Pixels corresponding to plastic get wrongly classified as wood on certain areas. This is caused by wood and plastic being relatively close to each other in the feature space as shown in Figure 11, and suggests the need for a more complex reflection model—thus requiring additional viewpoints—to distinguish materials with similar reflectance properties.

6 Conclusion and Outlook

An exploratory investigation and concept development for SG evaluation and MC using PCs and optical intensity values acquired with a state-of-the-art depth camera was carried out. The results for SG evaluation show that it is possible to classify geometry based on the level of geometric variability that it exhibits, by 3D feature analysis and supervised classification. The generated SG maps are easy to interpret, by both human and computer-aided analysis. Furthermore, the evaluation of the presented MC approach on a reduced subset of materials shows successful classification based on an approximated model of their

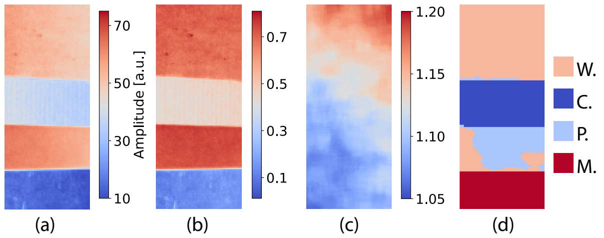

Figure 12. Result and intermediate steps for reflectance-based MC using 100 MHz data at distance of 1.5 m: a) measured intensities, b) corrected intensities, c) illumination distribution, and d) classified pixels, with material initials used for label names.

diffuse reflectance properties estimated from two viewpoints. The proposed method does not require in-situ calibration and can be easily extended to more complex reflectance models—hence potentially increasing classification robustness—by adding measurements from additional viewpoints.

The ultimate goal of the two proposed approaches is their in-line application within the robotic spraying process, providing feedback regarding geometric and material properties to adapt the process accordingly. On the one hand, the SG classes indicate whether further surface treatment in the fabrication process is needed or if the desired SG is met. On the other hand MC results can provide information on local coverage and volume of each sprayed material and also indicate early hydration states of the concrete to automatically regulate the time-window between consecutively sprayed layers.

As a continuation two main further investigations will be addressed. Firstly, the performance of the SG evaluation and MC approaches should be assessed on surfaces actually produced within the spraying process. Secondly, the combination of both approaches for enhanced performance should be explored. The latter means that geometry information (i.e. distances, orientation of the sample, etc.) should be used directly within the MC process where necessary, and vice-versa. Having MC results as one of the features within the SG evaluation, the material within the scene might contain information relevant for the SG classification.

Acknowledgments

This research was supported by the NCCR Digital Fabrication, funded by the Swiss National Science Foundation (NCCR Digital Fabrication Agreement #51NF40-182887).

The authors would like to thank Reckli GmBh for providing the moulds for concrete samples and to Andreas Reusser (NCCR dfab) for the help with concrete casting.

References

[1] Freek P Bos, Zeeshan Y Ahmed, Rob JM Wolfs, and Theo AM Salet. 3d printing concrete with reinforcement. In *High Tech Concrete: Where Technology and Engineering Meet*, pages 2484–2493. Springer, 2018.

[2] Rob Wolfs. *Experimental characterization and numerical modelling of 3D printed concrete: controlling structural behaviour in the fresh and hardened state*. PhD thesis, Technische Universiteit Eindhoven, Netherlands, 2019.

[3] Pablo Rodríguez-Gonzálvez and Gabriele Guidi. Rgb-d sensors data quality assessment and improvement for advanced applications. In *RGB-D Image Analysis and Processing*, pages 67–86. Springer, 2019.

[4] Benjamin Langmann, Klaus Hartmann, and Otmar Loffeld. Depth camera technology comparison and performance evaluation. In *ICPRAM (2)*, pages 438–444. Citeseer, 2012.

[5] Radu Horaud, Miles Hansard, Georgios Evangelidis, and Clément Ménier. An overview of depth cameras and range scanners based on time-of-flight technologies. *Machine vision and applications*, 27(7):1005–1020, 2016.

[6] *Time-of-flight forged ahead: design tips to boost 3D performance and cut integration time and cost*. Lucid Vision Labs, 2019. Manual, v.1.3, 18.10.2019.

[7] Monica Carfagni, Rocco Furferi, Lapo Governi, Michaela Servi, Francesca Uccheddu, and Yary Volpe. On the performance of the intel sr300 depth camera: metrological and critical characterization. *IEEE Sensors Journal*, 17(14):4508–4519, 2017.

[8] J Raja, B Muralikrishnan, and Shengyu Fu. Recent advances in separation of roughness, waviness and form. *Precision Engineering*, 26(2):222–235, 2002.

[9] Ali N Akansu, Wouter A Serdijn, and Ivan W Selesnick. Emerging applications of wavelets: A review. *Physical communication*, 3(1):1–18, 2010.

[10] Martin Weinmann, Boris Jutzi, Stefan Hinz, and Clément Mallet. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105:286–304, 2015.