Inline Monitoring of 3D Concrete Printing using Computer Vision
Rodrigo Rill-García, Eva Dokladalova, Petr Dokládal, Jean-François Caron, Romain Mesnil, Pierre Margerit, Malo Charrier

To cite this version:
Rodrigo Rill-García, Eva Dokladalova, Petr Dokládal, Jean-François Caron, Romain Mesnil, et al.. Inline Monitoring of 3D Concrete Printing using Computer Vision. 2022. hal-03788879

HAL Id: hal-03788879
https://hal-mines-paristech.archives-ouvertes.fr/hal-03788879
Preprint submitted on 27 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract
The detection of anomalies is at the basis of any 3D printing control. In this paper, we propose a methodology for detection of anomalies based on computer vision. This methodology is composed of three modules: 1) image acquisition, 2) interlayer line and layer segmentation and 3) characterization of the local geometry and texture of the layers and detection of anomalies. The image acquisition is performed with a camera fixed to the printing nozzle. The proposed layer segmentation method recognizes and locates the lines separating the printed layers (F-score = 91%). The third module – taking as input the segmentation and the original image – evaluates the geometry of the layers and the texture of the material. The results are used to detect geometry anomalies when the values are outside the expected range. The material texture is classified into four classes of quality (macro-averaged F-score = 94%). We present the results and show the suitability of our methodology for automatic detection and localization of anomalies on images acquired during a printing session.

Keywords: Automatic monitoring, 3D concrete printing, Image processing, Deep learning

1. INTRODUCTION
Construction based on additive manufacturing is now heading towards its maturity [1], with practical examples of 3D concrete printing (3DCP) at the building scale [2]. However, despite these high-profile applications, there is still work to be done on material development, construction strategies and the process control to make this technology reproducible and certifiable.

In principle, extrusion-based 3D printing is based on an established workflow: 1) design of the part, 2) definition of the print path and 3) extrusion of the material (printing) [3]. Nevertheless, many different process parameters influence the 3DCP extrusion e.g. the velocity of the extrusion nozzle, the distance between the nozzle and the surface of deposition, the dosing pump pressure, the extruded concrete’s rheological properties, etc. For this reason, various monitoring methods have been proposed at different stages: before printing (determining the feasibility of the process), during printing (process feedback), and at post-production stages (post-printing processing) [4].

The primary objective of inline monitoring (during printing) is to allow the operators to carry out corrective actions in case of anomalous behaviour. This principle can be extended to automatic control systems, as seen in the literature for manufacturing processes at different size scales. For example, in [5], a feedforward control system is tested for fused filament fabrication. In [6], a closed-loop control system is tested for 3DCP. In both cases, a vision system is used to measure the parameter of interest: the layer width, which can be controlled by adjusting the velocity of the extrusion nozzle.

Nonetheless, the layer width is not the only parameter that can be measured to be able to take corrective action and minimize defects. Some examples of geometrical deformations caused by process defects, that can be identified by inspecting the layers’ thickness, are described in [7]:

- **Excessive velocity** appears when the velocity of the extrusion nozzle exceeds the extrusion rate, producing discontinuous layers also called longitudinal tearing (Figure 1a).
- **Over-pressing** occurs when the pressing force resulting from the extruded layer exceeds the strength of the penultimate layer, resulting in staggered layer patterns with a loss of control of layer width and thickness (Figure 1b).
- **Flow-out** occurs when the yield stress of the material is not sufficient to hold its own weight. This introduces distance between the printed piece and the extrusion nozzle, producing a poor material deposition.
The measurement of the layers’ thickness can be performed by vision systems too. Furthermore, vision systems are not limited to the monitoring of geometrical properties, allowing the analysis of the observed surfaces in terms of properties such as the texture. In this work, using a lateral view of the printed piece, we propose a methodology for inline 3D concrete printing monitoring using computer vision methods (see Figure 2). Our contributions can be summarized as:

1. A method for automatic layer segmentation. We use a model based on Deep Learning (DL) to locate the interlayer lines in an image (i.e. the lines separating two printed layers), and we segment independent printed layers by using these lines as guide.

2. A monitoring method based on geometrical characterization of the layers. We use image processing to determine the local geometry, based on the interlayer lines, in terms of: orientation, curvature, layer thickness and distance to the nozzle. These measurements are used to detect and locate anomalies with respect to the expected values in the observed layers.

3. A monitoring method based on textural characterization of the layers. We use a machine learning approach to classify the extruded concrete, region-wise, based on textural properties. Textures are classified either as good or as one of fluid, dry and tearing. These classes, which are dependent on the water content of the mixture, are highly related to the layer’s rheological properties.

The rest of this paper is organized as follows. In section 2, we make a literature review of approaches used for inline 3DCP monitoring. In section 3, we describe our experimental setup, focusing on the inline image acquisition. In section 4, we discuss the properties of the images used in our experiments and the method used for interlayer line segmentation. The methods for layer geometry characterization and texture classification are described in section 5. In section 6, we show our results on experimental images acquired in a 3DCP laboratory. In section 7, we discuss these results and close the paper by proposing future improvements.

2. RELATED LITERATURE

For inline monitoring of 3DCP, a lot of research has been published on rheological properties of the material. For example, an online yield strength measurement on an uni-axial flow [8], where the mass of concrete drops is measured, provides information about the yield strength and heterogeneity of the material before the printing. The material properties can be controlled in real time using approaches such as a two-component strategy – with the addition of accelerator at the nozzle outlet [9]. This type of information is useful because insufficient stiffness or strength of fresh concrete can lead to an unacceptable cumulative error or even failure, as observed in [10]. On the other hand, other parameters may not make the printing process completely fail but would lead to the creation of visible defects leading to the rejection of the 3D printed part afterwards [11].
As a non-invasive approach to inline monitoring, methods based on specialized optical sensors have shown promising results to detect such defects. Most published research in this direction focuses on geometrical features. For example, a 1D Time-of-Flight distance sensor was mounted on the printing nozzle in [12]. The goal was to ensure that the printing nozzle is always at an adequate distance with respect to the last printed layer, so that the extruded material does not get unexpected deformations in its trajectory from the nozzle to the surface of deposition.

This method allows compensating the effect of non-planar print beds, and it forces the robot to extrude material at a proper distance even if lower layers are deformed by the weight of recent layers (flow-out). However, if such a flow-out deformation exist, the final result will be lower than the reference model, with excessively wide layers. This problem can be solved with approaches such as the one proposed by [13]. There, the printing method is shotcrete rather than layer extrusion; for printing monitoring, the authors use laser triangulation. Similarly to [12], the authors measure the distance from the nozzle to the last printed layer. In this work, however, the used printing method allows compensating the height of the printed piece if the bottom layers suffered deformation.

Although this method allows correcting the height of the piece, any anomaly in the width of the layer (over-pressing) is ignored. With respect to this problem, we can find alternatives like the one discussed in [14] based on computer vision.

2.1. Computer vision

3DCP monitoring can be performed with data retrieved from economic devices such as RGB cameras. For example, the work of [15] proposed a monitoring system fully based on computer vision, using a camera fixed to the extrusion nozzle. In that work, the image is first converted to grayscale and blurred with a Gaussian filter; the filtered image is binarized with the Otsu’s method to segment the printed layer from the background. Finally, the authors map the width of the segmented layer from pixels to inches (with a top view). The patent number US 8944799 B2 [16] was generated based on this method that adopts computer vision techniques for contour crafting. Low extrusion rates are expected to create thinner layers, while wider layers are produced by high extrusion rates. Another similar work is presented in [6], where the borders of the printed layer are detected as edges. There, the width of the layers is associated with the velocity of the extrusion nozzle, rather than the extrusion rate.

Towards more robust vision systems, DL has gained a significant attention in recent years. For printed concrete monitoring, we can see the example of [17]. There, DL-based segmentation is performed to distinguish a contour-crafting-printed piece from background (similarly to [15]). In that work, the camera acquires images from an external, post-printing, lateral view rather than from an inline top view. This allows simultaneously monitoring the piece’s height and the interlayer lines. Once the printed piece is isolated from the background in the captured images, a surface smoothing filter is used along with the Canny edge detector to identify the interlayer lines in the printed piece. Then, the Hough transform is used to estimate, region-wise, the orientation of these lines. Finally, a module for defect detection locates lines with unexpected orientations inside the printed piece.

For 3D geometry assessment, [18] compares the 3D model used for printing with a point cloud obtained (post-printing) by scanning the printed piece on a rotatory base. Using techniques from mathematical morphology, the authors calculate a distribution of cloud-to-cloud distances; this distribution allows quantifying global (all piece) and local (layer-wise) errors with respect to the model used to design the printing path.

As seen in the reviewed literature, optical sensors are useful for non-invasive 3DCP monitoring methods. However, research on these methods is still limited in the literature. Furthermore, research on methods based on computer vision is even more scarce. While geometrical assessment is a popular trend in methods using computer vision, monitoring of the surface has been ignored in 3DCP. In fact, surface analysis for inline monitoring of small-scale extrusion is an active topic of research [19]. Among the methods based on computer vision, the ones based on texture analysis for surface quality assessment can be extended for the monitoring of extruded concrete.

In this work, we propose to detect anomalies in both aspects – the geometry and the texture – with a single camera, during the printing process. Our methodology is compared in Table 1 with the different works based on optical sensors that were discussed in this section. With the experiments presented in this work, we will show that: 1) an analysis of the interlayer lines allows measuring multiple geometrical parameters of the printed piece to locate anomalies, making the segmentation of these lines an important task; 2) the texture of the extruded layers allows locating abnormal regions when the mixture presents anomalies in the water content; 3) the analysis of the geometry and the texture can be performed from the input of a single RGB camera.

3. EXPERIMENTAL SETUP

For the results presented in this paper, we created a dataset from images recorded during a printing session run in the 3D printing laboratory of Ecole des Ponts ParisTech. The cell is composed of a 6-axes industrial robot ABB IRB 6620\(^*\) using the xHEAD printing head developed by XTreee\(^*\)as extruder. It relies on a bi-component strategy, similar to the one described in [20], with an external feeding pump for the concrete, and a dosing pump for concrete.

Table 1: Summary of methods based on optics for 3DCP inline monitoring.

<table>
<thead>
<tr>
<th>Work</th>
<th>Sensor</th>
<th>Type of property</th>
<th>Measurements</th>
<th>Approximate cost&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>[12]</td>
<td>1D Time-of-Flight</td>
<td>Geometrical</td>
<td>Layer to nozzle height (top view)</td>
<td>5 USD</td>
</tr>
<tr>
<td>[13]</td>
<td>Laser triangulation</td>
<td>Geometrical</td>
<td>Layer to nozzle height (top view)</td>
<td>6500 USD</td>
</tr>
<tr>
<td>[15]</td>
<td>RGB camera</td>
<td>Geometrical</td>
<td>Layer width (top view)</td>
<td>40 USD</td>
</tr>
<tr>
<td>[17]</td>
<td>RGB camera</td>
<td>Geometrical</td>
<td>Layer orientation (lateral view)</td>
<td>40 USD</td>
</tr>
<tr>
<td>[18]</td>
<td>3D scanner</td>
<td>Geometrical</td>
<td>Printed piece cloud to 3D model cloud distance</td>
<td>472 USD</td>
</tr>
<tr>
<td>Ours</td>
<td>RGB camera</td>
<td>Geometrical, textural</td>
<td>Layer to nozzle height; layer thickness; interlayer line orientation and curvature; texture classification (lateral view)</td>
<td>57 USD</td>
</tr>
</tbody>
</table>

<sup>a</sup> Estimated with base on the sensor models and/or specifications reported in the respective cited works.

and accelerator in the printing head. A mixing paddle guarantees the homogeneity of dispersion of accelerator in a mixing chamber. The material is the NAG3 concrete developed by Lafarge and tested in [21]. The nominal speed is 80mm/s, with a nozzle diameter of 20mm, while the pipe diameter is 30mm. The typical flow rate is 1L/min, but may slightly vary during the experiment.

The robot trajectories are constructed from a 3D model and a simple slicing algorithm in Grasshopper [22], which outputs a series of planes (targets). These targets are later converted into RobotStudio files through HAL [23], a general programming language for industrial robotics. Constraints of orientation were added in order to guarantee the perpendicularity between the camera frame and the deposited layer; the inverse kinematics is solved with HAL, which also indicates if a position is not accessible and allows either repositioning the printed piece in a more accessible region or changing its geometry in the 3D modelling environment.

The robot controller controls the position and speed of the robot, the flow rate of the dosing pump and the flow rate of accelerator. The printing parameters used during the session were dynamically varied to purposely create geometrical and textural imperfections (see Figure 3).

Using this setup, the proposed monitoring methodology consists of 3 sequential modules (see blue dashed rect-

![Figure 3](image_url)

**Figure 3:** a) Experimental sample printed with purposely created defects. b) Example of acceptable printing without purposely created defects.

angles in Figure 2:

1. Image acquisition
2. Interlayer line and layer segmentation
3. Local layer characterization and anomaly detection
4. IMAGE ACQUISITION AND INTERLAYER LINE SEGMENTATION

Several setups for image acquisition are possible. As illustrated in Figure 4a, we chose a camera fixed to the nozzle in order to have a fixed point of view with respect to the extrusion zone.

4.1. Image acquisition setup

To avoid blur in the acquired images, caused by regions of the printed piece outside the depth of field of the camera, the camera orientation should remain perpendicular with respect to the printed wall. This normally requires a seven axis in the robot that allows the extrusion nozzle to rotate around its longitudinal axis. This problem of orientation also exists when printing with non-circular nozzles, which are common place in 3DCP; although out of the scope of this paper, there are technical solutions to this issue [24].

Regarding image resolution, it should allow correctly capturing texture i.e. identifying the aggregates in the concrete. In the experiments for this article, the largest aggregate is roughly 0.3mm. The desired camera resolution is therefore >3.3 px/mm. Due to eventual deformations or lack of control of width, small apertures are preferred in our applications: for a smaller aperture, the depth of field increases, allowing obtaining sharp images even if the distance between the observed concrete and the lens differs from the expected value.

Small aperture is usually used with long exposure times, which implies a trade-off on this parameter, because mitigation of motion blur requires short exposure times. Here, we are mostly in the case of linear motion blur. If we want the blur to be limited to 1 pixel, we get a constraint on exposure time. Using experimental data, we get an upper bound of $t_{\text{exposure}} \sim 1/3000s$, which holds for a wide range of printed concretes, where typical aggregate size remains submillimetric and robot speed in the 100mm/s range.

4.1.1. Properties of the acquired images

The distance of the lens was adjusted to contain approximately 4 layers per frame, with an expected layer thickness of 6mm. To acquire images, a Raspberry Pi Camera High Quality Module featuring a SONY IMX477 CMOS sensor with a macro lens was fixed on the printing nozzle (more camera specifications available at [25]). This allowed having a fixed distance (~13cm) between the lens and the extruded material. Additionally, we mounted a led lamp on the nozzle to allow a short, and constant, exposure time; this is illustrated in Figure 4a. The image resolution is 1280 x 960 pixels; therefore, our images contain \(\approx 40\) px/mm. The focus, exposure time and lens aperture were manually adjusted prior to printing the piece and fixed during extrusion.

To have uniformity in our final dataset, from the acquired images we chose a subset of pictures meeting 2 conditions: 1) no background is visible at the sides of the printed piece, 2) the center of the nozzle is coplanar with the center of the observed wall (as illustrated in Figure 4a). We preserved 628 raw images in total; a schematic diagram of a typical frame with acceptable printing is shown in Figure 4b.

4.2. Interlayer line and layer segmentation

Once an image is acquired, the next step in our methodology is to segment its interlayer lines. On dry pieces, the detection of these lines can be based on finding dark edges (see Figure 5a). However, fresh concrete exhibits specular reflectance as well as layer merging and superposition. Consequently, interlayer lines can take different aspects: black, bright or even not visible at all (see Figure 5b).

Furthermore, given the constraints imposed by inline image acquisition (as described in subsection 4.1), the problem becomes particularly hard: loss of focus, unexpected motion blur, camera vibration, polluted air (presence of particles), etc.

To approach this problem, we trained a neural network originally used for crack segmentation: U-VGG19 [26]. It is a fully convolutional neural network inspired by U-net which uses VGG19 as backbone model; the architecture of this encoder-decoder network is illustrated in Figure 6. By using this network, we get as output a binary mask with the location of the interlayer lines in the input image (as illustrated in Figure 7).

Given an input image $I$, let us denote by $S \subset \mathbb{Z}^2$ the interlayer lines segmented by the model. This segmentation $S$ can be used to extract the layers delimited vertically.
between each two lines. Let us denote by \( \text{ROI} \subset \mathbb{Z}^2 \) this set of layers (see Figure 7c).

5. **INLINE 3D CONCRETE PRINTING MONITORING**

Once the interlayer lines and the independent layers are segmented, the next step in our methodology is to perform a local layer characterization for anomaly detection.

5.1. **Geometrical characterization**

The proposed geometrical characterization consists of local measures of the geometry of the printed layers and interlayer lines. These measures, and the posterior anomaly detection, are performed using the following methods.

5.1.1. **Local orientation of interlayer line**

As discussed by \cite{17}, studying the orientation of the interlayer lines allows detecting deformations in the printed layers. To do this, we propose a method based on image filtering using rotating structuring elements \cite{27}.

First, we perform a thinning \cite{28} to approximate the interlayer lines to 1-pixel-width lines. Let \( S_T \subset \mathbb{Z}^2 \) denote the thinned image, \( R(d, l) \) a line oriented by \( d \) with length \( l \), and \( D \) the set of possible orientations \( d \in [-90, 90) \). The line length \( l \) is chosen to be 1/5 of the image width. This hyperparameter presents a trade off between the available line angles \cite{27} and the locality of the measurement. The value (256 pixels) was selected to be similar to the expected layer thickness, but it is not critical for the method. Let \( p \in \mathbb{Z}^2 \) be a point in an image. Then, \( F(p, d) = [S_T * R(d, l)](p) \) is the convolution of \( S_T \) with a line oriented in \( d \). Finally, our local orientation map \( \phi \) is defined as:

\[
\phi(p) = \begin{cases} 
\arg \max_{d \in D} F(p, d), & p \in S_T \\
\text{undefined}, & \text{otherwise}
\end{cases}
\] (1)

5.1.2. **Local curvature of interlayer line**

Let \( \hat{L}_n \) be a piece-wise approximation of the \( n \)-th interlayer line in \( S_T \), using \( m \) cubic splines \cite{29}:

\[
\hat{L}_n(t) = \begin{cases} 
s_n^i(t - t_0), & t_0 \leq t < t_1 \\
\vdots \\
s_n^m(t - t_{m-1}), & t_{m-1} \leq t < t_m
\end{cases}
\] (2)

The knots are selected to split the interlayer line of interest into \( m \) fixed-length segments. This length is a hyperparameter, and it implies a trade-off between the accuracy and the smoothness of the spline approximation. The value was chosen to be 1/10 of the image width (128 pixels), but it is not critical for the method. The independent variable \( t \) corresponds to a parametric representation of the spline \( s_n^i(t) = (x(t), y(t)) \).

Since this smooth approximation is twice differentiable, the local curvature \( \kappa \) of \( \hat{L}_n \) at point \( t \) can be computed as:
5.1.3. Local thickness of layer

The local thickness of a layer is determined by the distance between its two interlayer lines. We calculate this thickness using the fast Euclidean Distance Transform (EDT) \[30\].

The function \(D : \mathbb{Z}^2 \rightarrow \mathbb{R}^+\) is the result of applying the EDT to \(S_T\) (see Figure 8a). The local maxima of \(D\) indicate the centers of the printed layers (see Figure 8(b-c)). Since we expect the layers to be nearly horizontal in the input picture, we estimate the locations of these local maxima along the vertical direction. Let \(k_1 = (0, -1, 1)^T\) and \(k_2 = (1, -1, 0)^T\) be vertical kernels. The convolution of the distance map \(D\) with these kernels allows detecting the maxima of \(D\):

\[
\kappa(t) = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{\frac{3}{2}}}
\]

The primes refer to the derivatives with respect to \(t\). The magnitude of \(\kappa(t)\) is the inverse of the radius of the osculating circle touching the point \((x, y)\) defined by \(L_n(t)\). The sign indicates the concavity.

### 5.1.4. Relative nozzle height

Similarly to \[12, 13\], we measure the relative height of the printing nozzle with respect to the surface of deposition. From the lateral point of view proposed in this paper, this height corresponds to the vertical distance from the nozzle to the line receiving the currently printed layer.

Since the camera is fixed to the robot, the center of the nozzle is at a constant position \(c = (x_c, y_c)\) in the image. Let \(L_2\), the second interlayer line from top to bottom in the image, be the set of points \(p = (x_p, y_p)\) separating the currently printed layer from the layer of deposition (see Figure 9). Then, the relative nozzle height is:

\[
H(p) = \begin{cases} 
 y_p - y_c, & p \in L_2 \\
 \text{undefined}, & \text{otherwise} 
\end{cases}
\]  

With the vertical axis pointing downwards in an image, this height is positive as long as \(p\) is below the nozzle.

### 5.1.5. Geometrical anomalies

To summarize the information retrieved by Eq. \[4\], \[8\], \[5\] and \[6\] (orientation, curvature, thickness and height, respectively), we obtain a histogram per measure. Under good printing conditions, the distributions generated from these histograms should be centered and concentrated inside a range of admissible values.

The principle is illustrated in Figure 10. There, we show the interlayer segmentation maps from a frame with acceptable printing and one exhibiting signs of over-squeezing (see Figure 1 for reference). Additionally, we show the distributions of layer thickness measured pixel-wise by our method for each frame. The hypothesis is that a defect can be modeled as a deviation from the distribution expected for a defect-less printing (Figure 10c). Then it becomes straightforward to identify the position of these deviations and localize the corresponding portion of layer in the frame. Since the fixed pose of the camera with respect to the nozzle is known, and the pose of the nozzle at the time of acquiring the analyzed frame can be retrieved from the robot’s control software, it is possible to calculate the physical localization of the anomalies in the printed piece.

In the next section, we present the characterization of the texture of the printed layers.

5.2. Texture characterization

The rheological properties of the extruded material are directly related to the composition of the mixture – which
impacts the structural properties of the printed piece. In this work, we focus on the water content: lack of water can cause cold joints at the interface, but an excess of water makes the layers more prone to deformation or even collapse. The water content is one of the several factors affecting the texture of the extruded layers (see Figure 11). When the texture exhibits abnormal properties, it becomes a good visual indicator of anomalies in the printing process.

In our method, the texture of the printed piece is analysed locally, layer by layer, in small windows as shown in Figure 12. The height of the windows is adapted as to cover the entire thickness of either layer, excluding adjacent layers. The windows have a fixed width of 200 pixels (~5mm) and are adjacent each to another. The value of this hyperparameter (chosen to have near-square windows) presents a trade off between the available area to analyze per window and the locality of the analysis. Each window is then independently classified, either as good or one anomalous class.

5.2.1. Pre-processing

Before extracting the windows, we level the image’s grayscale intensities to preserve textural information while reducing the effects of lightning and shadows. Given the input image $I$, first we obtain a local-mean-intensity im-
figure using a Gaussian filter with $\sigma = 40$. By subtracting the filtered image from $I$, we obtain a new image with an intensity distribution centered around 0. We further subtract the minimum to avoid negative values. The result of this whole process is illustrated in Figure 13.

5.2.2. Texture features

Each window is analyzed individually to provide a label using a machine learning approach. Our feature extraction is based on two visual descriptors: gray-level co-occurrence matrices (GLCMs) and local binary patterns (LBP). These descriptors are obtained from the maximum rectangular box contained within concrete pixels in the window. Let each of these boxes be called a $T_{box}$.

Before feature extraction, each $T_{box}$ is quantized to $q$ discrete values. To avoid that outlier intensity values have an undesired influence, each $T_{box}$ is clipped to values in the range $\text{mean}(T_{box}) \pm 3 \cdot \text{std}(T_{box})$ before quantization. Let the quantized version of the image, in the range [0, $q-1$], be called a $T_{q}$. box.

GLCMs. Let define a GLCM \cite{31} as:

$$GLCM_{\Delta x, \Delta y}(i, j) = \sum_{x=1}^{h} \sum_{y=1}^{w} \begin{cases} 1, & \text{if } I(x, y)=i \text{ and } I(x + \Delta x, y + \Delta y)=j \\ 0, & \text{otherwise} \end{cases}$$

(7)

Here, $h$ and $w$ are the height and the width of an analyzed image $I$, respectively; $\Delta x$ and $\Delta y$ are a horizontal and a vertical offset, respectively. $I(x, y)$ returns the intensity value of the pixel at the position $x, y$ in $I$.

GLCMs are distributions of co-occurring pixel values at a given scale (the chosen offset). From preliminary experiments, $q = 12$ was chosen for GLCM calculation. With 12 intensity levels in a $T_{q}^{box}$, per each offset we obtain a square 12x12 matrix (see Figure 14). Per each GLCM, we obtain scalar features in terms of statistical properties. With $P_{i,j}$ the value of the normalized GLCM at the position $i, j$ (i.e. the probability of the intensity pair $(i,j)$), the features are obtained based on 5 properties:

Contrast: the expected squared difference of intensities in the GLCM. It is defined as:

$$\sum_{i,j=0}^{\text{levels}-1} P_{i,j} \cdot (i-j)^2$$

(8)

Dissimilarity: the expected absolute difference of intensities in the GLCM. It is defined as:

$$\sum_{i,j=0}^{\text{levels}-1} P_{i,j} \cdot |i-j|$$

(9)

Homogeneity: a measure of the closeness of the distribution of elements in the GLCM to its diagonal. It is defined as:

$$\sum_{i,j=0}^{\text{levels}-1} \frac{P_{i,j}}{1 + (i-j)^2} \cdot |i-j|$$

(10)

Energy: the square root of the sum of squared elements in the GLCM. It is defined as:

$$\sqrt{\sum_{i,j=0}^{\text{levels}-1} P_{i,j}^2}$$

(11)

Correlation: a measure of how correlated are the pixels to their neighbors with the given offset. It is defined as:

$$\sum_{i,j=0}^{\text{levels}-1} P_{i,j} \cdot \frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2 \cdot \sigma_j^2}}$$

(12)

Due to the printing process, the texture properties are anisotropic. Because of this, we extract GLCMs with offsets in the horizontal and vertical directions. Per each direction, we use offset distances from 1 to 50 with step 1. The total number of features obtained per $T_{box}$ is 500 (2 directions x 50 distances x 5 statistical properties).

LBP. This descriptor consists of assigning a $n$-bit binary number to each pixel depending on its $n$ neighbors...
with radius \( R \). Following the neighbors along the hypothetical circle of radius \( R \), each bit is assigned a 0 if the center pixel’s intensity value is greater than the corresponding neighbor’s value. Otherwise, the bit is assigned a 1. The transformed image contains, per pixel, the decimal equivalent of the binary number calculated according to their neighbors. We use 8 neighbors, therefore the transformed images contain values in the range \([0, 255]\).

From preliminary experiments, \( q = 64 \) was chosen for quantization to extract LBP (see Figure 15). After getting the LBP transform of a \( T^q_{box} \), we calculate the histogram of intensities of the transformed image to extract a fixed-size vector per \( T^q_{box} \). To reduce the dimensionality of this vector, the histograms are calculated using 8 bins. These histograms are normalized so that their magnitudes are not influenced by the size of the \( T^q_{box} \).

Similarly to GLCMs, we use different radii to extract information at different scales. These radii are from 1 to 31 with step 10. The total number of features per \( T^q_{box} \) is 32 (4 radii \( \times \) 8 bins).

5.2.3. Classification

With the extracted textural features, we classify each \( T^q_{box} \) into 1 of 4 classes as illustrated in Figure 16: excessively fluid, good quality, excessively dry, and tearing.

The fluid class is characterized by a smooth surface. Since the fluidness is associated with high water content, this class is also characterized by a high specular reflectance. The good class corresponds to the desired material properties. The amount of present water is adequate, producing the appearance of more visible grains at the surface with higher homogeneity and a lower reflectance. The dry class is characterized by a rougher texture, caused by an increased amount of visible grains with bigger size. Since the amount of water is reduced, the material exhibits very low to no reflectance. However, this class is also determined by the absence of cracks. The tearing class is the one characterized by the appearance of cracks. It is also identified by the lack of reflectance and often by a rougher texture than the dry class.

From the images acquired during the printing process, we extracted and labeled a total of 111 texture windows. Each of these windows is labeled by the agreement of two annotators. The distribution of classes is: 24 fluid, 27 good, 24 dry and 36 tearing. The base dataset is composed by the 532 features calculated per each of these images (GLCM + LBP features), and their corresponding class labels. To avoid that some features dominate the classifier because of their magnitude, we standardize each one of them during training.

The model used to learn from this domain is a small convolutional neural network (CNN) with one convolutional input layer and one fully-connected output layer (see Figure 17). Because of the feature extraction implementation, adjacent elements in the feature vector represent the same property but with different offsets (see Figure 18). Since these features have a high likelihood of being correlated, we use the convolution to capture meaningful information from the neighborhood. The kernel size is 6, with a ReLu activation to obtain non-linear...
outputs. To reduce the likelihood of overfitting, we use dropout regularization with 30% probability. The output layer has 4 neurons and uses a softmax activation so that \( \sum_{c \in \text{classes}} P(c|\text{features}) = 1 \).

5.2.4. Textural anomalies

Similarly to the approach used for geometrical anomaly detection, we can summarize an image in terms of the histogram of window labels. In this case, the categorical distribution should be centered and concentrated in the good class. High occurrence of other classes indicates some anomaly in the printing process.

6. EXPERIMENT RESULTS

6.1. Segmentation

Once an image from the printing process is acquired, the next step is to segment the interlayer lines to obtain the binary segmentation maps.

We used 128 manually annotated images for training and 32 for testing. The labeling was done in Krita 4.4.5 using a circular brush with a diameter of 20 pixels (see Figure 19 for an example of manual segmentation). During training, we provided 256 × 256 image crops with a batch size of 4. We implemented U-VGG19 using Tensorflow 2.1.0. Similarly to [26], we initialized the weights of the encoder with those of the convolutional layers of VGG19 (pre-trained on ImageNet). We used the same optimizer and loss function i.e. Adam and a weighted combination of binary cross-entropy loss and dice score loss. The model is trained during 200 epochs, with an early stop if the testing loss didn’t improve for 20 consecutive epochs. The learning rate was decreased by a factor of 2 on testing loss plateau with 5 epochs tolerance; the initial learning rate was \( 10^{-4} \).

We evaluate the performance of U-VGG19 on the test images in terms of precision, recall and F-score. We calculate each score per image and we report the average over all the test images. There is a possibility of offsets between prediction and annotation caused by the inaccuracy of the annotators at the moment of locating the center of the interlayer lines. To address this, we use a tolerance margin similarly to works such as [33, 34]. The chosen tolerance is 2 pixels, meaning that predicted positive pixels no more than 2 pixels away from a positive pixel in the manual annotation are considered true positives. This represents approximately 0.5mm in real life. The scores are shown in Table 2.

<table>
<thead>
<tr>
<th>F-score (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>90.63</td>
<td>92.14</td>
<td>89.22</td>
</tr>
</tbody>
</table>

The average precision and the average recall have very similar values (around 90%). Therefore, U-VGG19 exhibits a balanced ability to detect interlayer lines while discriminating false positives. An example of automatic interlayer line segmentation using U-VGG19 is illustrated in Figure 19. With an average F-score of 90.63%, we use the predictions of U-VGG19 to perform the characterization methods proposed in this paper.

6.2. Texture classification

In this multi-class context, we evaluate the proposed method with 5-folds stratified cross-validation i.e. each fold has the same proportion of observations with a given class. To further reduce the likelihood of overfitting, per iteration we perform another stratified split to divide the training folds into a training and a validation split; the size of the validation split is half the size of the test split. Finally, we perform data augmentation on the images from the resulting training split, obtaining 16 additional images per texture window. The transformations are rotations (with angles -4, -3, -2, 2, 3, 4 degrees), zooms (with scale factors 0.97, 0.98, 0.99, 1.01, 1.02, 1.03) and illumination rescales and shifts defined by \( \alpha \cdot \text{intensity} + \beta \) (with pairs \( \alpha/\beta \): 0.9/0, 1.1/0, 1.0/-50, 1.0/50).
Per iteration, the approximate number of samples per split is 1326 training, 11 validation and 22 test. Validation and testing are performed on real images only, while the training split does not contain modified versions of the images in the validation and test splits. The approximate class distribution is 22% fluid, 24% good, 22% dry and 32% tearing. We train the CNN with a binary cross-entropy loss using the Adam optimizer with default parameters in Tensorflow 2.1.0. We train for a maximum of 2000 epochs with batch size 32, with an early stop if the loss in the validation split didn’t improve for 100 consecutive epochs.

We evaluate the trained models in terms of the macro-averaged F-score in the test splits; the macro F-score gives equal importance to all the classes. Table 3 shows the macro F-score per fold and per class; the single F-score per fold and class are shown too.

Table 3: F-scores(%) of texture classification using 5-folds stratified cross-validation.

<table>
<thead>
<tr>
<th>Class</th>
<th>Fold 1</th>
<th>Fold 2</th>
<th>Fold 3</th>
<th>Fold 4</th>
<th>Fold 5</th>
<th>Class average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluid</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Good</td>
<td>100</td>
<td>100</td>
<td>88.9</td>
<td>90.9</td>
<td>88.9</td>
<td>93.7</td>
</tr>
<tr>
<td>Dry</td>
<td>90.9</td>
<td>100</td>
<td>83.3</td>
<td>100</td>
<td>76.9</td>
<td>90.2</td>
</tr>
<tr>
<td>Tearing</td>
<td>93.3</td>
<td>100</td>
<td>92.3</td>
<td>92.3</td>
<td>83.3</td>
<td>92.2</td>
</tr>
<tr>
<td>Class macro-average</td>
<td>96.1</td>
<td>100</td>
<td>91.1</td>
<td>95.8</td>
<td>87.3</td>
<td></td>
</tr>
</tbody>
</table>

Among the four classes, fluid exhibited the best scores (100% in all folds). On the other side, the most difficult class was dry, with a 90.2% average over the 5 folds. The good and tearing classes have slightly better averages: 93.7 and 92.2%, respectively. From an analysis of the resulting confusion matrices per fold (see Figure 20), we discovered that the main source of these errors was classifying good and tearing images as dry. Nonetheless, this kind of confusion is understandable since the dry class is, physically, a transitory state between good printing and tearing. For curvature, we expect near-to-horizontal layers; the user-defined range is (-0.05, 0.05) degrees. For curvature, we expect near-tostraight layers; the range is (-0.05, 0.05) mm$^{-1}$. The expected layer thickness is 6mm; the range is (4.5, 7.5) mm. The nozzle should keep a height close to the expected layer thickness; the range is (5, 7) mm. Regarding texture, any class but good is considered as anomaly; the texture classes are associated to a color code similar to the color maps used for geometrical measurements: fluid/blue, good/green, dry/yellow, tearing/red. Additionally to the class, we show the probability of the prediction according to the classifier.

The first study case, analyzed in Figures 21 and 22, portrays an overall acceptable printing. As observed in Figure 21(b-e), there are almost no geometrical anomalies (except a few high curvature segments); consequently, the distributions obtained from these plots are contained inside the ranges of admissible values, as seen in Figure 22(b-e).

The second study case, analyzed in Figures 23 and 24, depicts a scenario with severe anomalies. As observed in Figure 24(b-e), the distributions of the geometrical measurements are very different from the ones in Figure 22 in all the distributions, there are high density regions outside the range of admissible values. The most extreme case is present in Figure 24e where the distribution is concentrated outside the admissible range. This behavior is likely to cause a coiling effect on the material deposition; in fact, as depicted in Figures 23c and 24c, we see high curvatures that can be directly related to this phenomenon.

Furthermore, with respect to the texture analysis, we see that the second layer from top to bottom exhibits many regions with tearing. This is the extreme case of texture anomalies, since it implies the appearance of cracks that can jeopardize the structural safety of the printed piece.

As shown by this study case, as well as the first one with overall acceptable printing, the proposed methodology is able to provide an inline characterization of the process based on visual inspection of the last printed layers. With this characterization, the proposed methodology detects and locates anomalies in the extruded layers.

7. CONCLUSIONS AND FUTURE WORK

In the present article, we proposed a methodology for computer-vision based, inline monitoring of 3D concrete printing. This methodology consists of three sequential modules: 1) image acquisition, 2) interlayer line and
layer segmentation, and 3) local layer characterization and anomaly detection.

With the presented experiments using the aforementioned methodology, we show that:

- The use of Deep Learning allows an inline segmentation of the interlayer lines separating adjacent extruded layers. The proposed model obtains an F-score of 91%. We demonstrate that locating these lines in the analyzed images allows a further segmentation of the independent layers.

- An analysis of the segmented interlayer lines allows measuring multiple geometrical properties of the printed piece simultaneously. We demonstrate that the measurement of these properties allows detecting and locating anomalies in the printed piece.

- An analysis of the texture of the independent layers allows detecting when the extruded material exhibits anomalies in the water content. We suggest 4 classes for this analysis, and the model proposed to classify the textures obtains a macro-averaged F-score of 94%.

Our experiments show that the geometrical and the textural analyses can be performed simultaneously from the input of a single RGB camera. With this, we extend the catalogue of possible geometrical measurements using computer vision with respect to the current proposals in the literature. Additionally, we are the first to propose an analysis based on texture for 3D concrete printing, inspired by the approaches already used in small-scale, extruded-based additive manufacturing.

In the two presented study cases, we show that our methodology provides the operator with the exact position and nature of the detected anomalies. If the results
Figure 22: Distributions from the first study case. The dotted lines represent the range of admissible values; regions with values outside these ranges are detected as anomalies.

(a) Image to analyze as captured by the camera
(b) Local orientations of the interlayer lines
(c) Local curvatures of the interlayer lines
(d) Local thickness of the layers. The thickness is associated to the layer centers (in color); the white lines represent the interlayer lines
(e) Local heights of the nozzle with respect to the last printed layer
(f) Local textures. The white lines are the borders of the independently analyzed regions; the text shows the assigned class and probability

Figure 23: Plots from the second study case. The black pixels in the plots correspond to undefined values in the 2D space.
are recorded, they can serve as a quality report after the printing process is finished. This report can serve as a proof of acceptability of the printed piece. Inversely, it can also serve, when severe anomalies are detected to reject the piece – and even before the end the printing to avoid material and time waste. Finally, the nature and severity of the detected anomalies can be reported to the operator to make corrective adjustments.

A future research could extend this monitoring towards a closed-loop control. The nature and the severity of the anomalies could be used as feedback to an automated decision system issuing a corrective action on the printing system. It could stem from a rule-based expert system, vector-to-action dictionaries, supervised machine learning models, etc.

The images and annotations used for the results presented in this paper are available as a public dataset named I3DCP: https://github.com/Sutadasuto/I3DCP

ACKNOWLEDGEMENTS

This work has been supported by the project DiXite. Initiated in 2018, DiXite (Digital Construction Site) is a project of the I-SITE FUTURE, a French initiative to answer the challenges of sustainable city. This research project was financially supported by I-SITE FUTURE (ANR-16-IDEX-0003).

References

[2] All3DP Pro, 8 Biggest Companies Building 3D Printed Houses All3DPro ProAvailable at https://all3dp.com/2/best-companies-building-3d-printed-houses/