Improvement of Geometric Quality Inspection and Process Efficiency in Additive Manufacturing

Yu Jin
University of Arkansas, Fayetteville

Follow this and additional works at: https://scholarworks.uark.edu/etd

Part of the Industrial Technology Commons, Operational Research Commons, and the Systems Engineering Commons

Citation

This Dissertation is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact ccmiddle@uark.edu.
Improvement of Geometric Quality Inspection and Process Efficiency in Additive Manufacturing

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Engineering with a concentration in Industrial Engineering

by

Yu (Chelsea) Jin
Jinan University
Bachelor of Science in Network Engineering, 2014
University of Michigan
Master of Science in Manufacturing Engineering, 2015

May 2020
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Harry Pierson, Ph.D.
Dissertation Co-Director

Haitao Liao, Ph.D.
Dissertation Co-Director

Sarah Nurre Pinkley, Ph.D.
Committee Member

Zhenghui Sha, Ph.D.
Committee Member
Abstract

Additive manufacturing (AM) has been known for its ability of producing complex geometries in flexible production environments. In recent decades, it has attracted increasing attention and interest of different industrial sectors. However, there are still some technical challenges hindering the wide application of AM. One major barrier is the limited dimensional accuracy of AM produced parts, especially for industrial sectors such as aerospace and biomedical engineering, where high geometric accuracy is required. Nevertheless, traditional quality inspection techniques might not perform well due to the complexity and flexibility of AM fabricated parts. Another issue, which is brought up from the growing demand for large-scale 3D printing in these industry sectors, is the limited fabrication speed of AM processes. However, how to improve the fabrication efficiency without sacrificing the geometric quality is still a challenging problem that has not been well addressed. In this work, new geometric inspection methods are proposed for both offline and online inspection paradigms, and a layer-by-layer toolpath optimization model is proposed to further improve the fabrication efficiency of AM processes without degrading the resolution. First, a novel Location-Orientation-Shape (LOS) distribution derived from 3D scanning output is proposed to improve the offline inspection in detecting and distinguishing positional and dimensional non-conformities of features. Second, the online geometric inspection is improved by a multi-resolution alignment and inspection framework based on wavelet decomposition and design of experiments (DOE). The new framework is able to improve the alignment accuracy and to distinguish different sources of error based on the shape deviation of each layer. In addition, a quickest change point detection method is used to identify the layer where the earliest change of systematic deviation distribution occurs during the printing process. Third, to further improve the printing efficiency without sacrificing the quality of each layer, a toolpath allocation and scheduling optimization model is proposed based on a concurrent AM process that allows multiple extruders to work collaboratively on the same layer. For each perspective of improvements, numerical studies are provided to emphasize the theoretical and practical meanings of proposed methodologies.
## Contents

1 Introduction 1

2 Scale and Pose Invariant Feature-based Quality Inspection 4  
   2.1 Introduction ......................................................... 4  
   2.2 Literature Review .................................................. 8  
   2.3 Theory ................................................................. 11  
   2.4 Methods .............................................................. 13  
      2.4.1 Inspection based on LOS Distribution ......................... 13  
      2.4.2 Point Cloud Downsampling .................................. 18  
      2.4.3 Experimental Description ................................... 19  
   2.5 Results .............................................................. 20  
      2.5.1 Experiment 1 .................................................... 20  
      2.5.2 Experiment 2 .................................................... 23  
      2.5.3 Experiment 3 .................................................... 29  
      2.5.4 Experiment 4 .................................................... 32  
   2.6 Discussion ......................................................... 33  
   2.7 Conclusion .......................................................... 36  
Reference ................................................................. 37

3 Multi-resolution In-plane Alignment and Error Quantification 42  
   3.1 Introduction ......................................................... 42  
   3.2 Literature Review .................................................. 46  
      3.2.1 Point Cloud Alignment ....................................... 46  
      3.2.2 Deviation Modeling and Quantification in AM processes ... 47  
   3.3 Method .............................................................. 49  
      3.3.1 Shape Profile .................................................... 50  
      3.3.2 Wavelet Decomposition ...................................... 51
4.5 Conclusion ......................................................... 109
Reference .......................................................... 110

5 Summary ......................................................... 114
List of Figures

1.1 An overview of three research problems ................................. 3

2.1 (a) Topology-optimized mounting bracket for an aerospace application (Image courtesy of EOS GmbH); (b) A freeform design lacking parametric surfaces for establishing datums and fixturing (Image courtesy of EOS North America) ........ 6

2.2 Example of misalignment using the ICP method. ...................... 7

2.3 LOS inspection process overview. ........................................ 14

2.4 Experimental Geometries .................................................. 19

2.5 The point cloud associated with test object 1 was modified for scale, orientation, and point cloud density to test the sensitivity of the LOS distribution to these factors. 21

2.6 LOS distributions comparison for test object 1 are invariant to pose and scale, but are sensitive to point cloud density. ........................................ 22

2.7 Objects printed for experiment 2(a-c); features of interest for experiment 2. ....... 23

2.8 Shape inspection results for test objects 2a and 2b detect the shape defect introduced in test object 2b. ........................................ 24

2.9 Datum feature and four target features. ................................... 25

2.10 LOS inspection result for test objects 2a and 2c detect the pose defect introduced in test object 2c. ........................................ 26

2.11 The LOS distribution comparison for test object 2a indicates no features out of place with respect to each other. ........................................ 27

2.12 The LOS distribution comparison for test object 2c indicates pose problems with feature 3 relative to features 1, 2, and 4. .................. 27

2.13 An example when feature indexing of nominal model and sample parts is inconsistent. ........................................ 28

2.14 Feature re-indexing was successful for test object 2a. .................. 28

2.15 Feature re-indexing was successful for test object 2c despite the pose error associated with feature #3. ........................................ 29
2.16 Test objects for Experiment 3; the shape distribution of each feature is compared in (c). ................................................................. 30
2.17 Feature re-indexing was successful for test object 3. ......................... 31
2.18 Three test objects were printed for Experiment 4: one unaltered and two with pose errors introduced. ......................................... 32
2.19 LOS inspection results for Experiment 4 indicate good shape fidelity for all test objects, but clearly differentiate the pose errors that exist in test objects 4b and 4c. 33
3.1 An overview of the automatic multi-resolution alignment and inspection framework. ................................................................. 49
3.2 An example of shape representation in CCS, PCS, and shape profile. .......... 51
3.3 Wavelet decomposition of shape profile. ............................................ 52
3.4 An example of finding best matches for the 2nd level of the sample coefficient profile in the extended reference coefficient profile. ..................... 54
3.5 Find the optimal phase shift for the 2nd level of the coefficient profile. ...... 55
3.6 Multi-resolution alignment of detail coefficient profiles and shape profiles. .... 55
3.7 Demonstration of change point in a time series. ................................. 59
3.8 Overview of simulation data. ............................................................ 61
3.9 The QCD result based on single post-change model. ............................ 61
3.10 The QCD result based on multiple post-change model. ......................... 62
3.11 The plot of ADD under different variance. ........................................ 63
3.12 Performance of QCD under different post-change distribution $N(\mu, \sigma^2)$. .......... 63
3.13 Experiment setup. ................................................................. 64
3.14 Three sample layers. .................................................................. 65
3.15 Shape profile of the 1st layer. ......................................................... 65
3.16 Optimization of alignment parameters for the 1st layer. ....................... 66
3.17 4-level multi-resolution alignment results. ....................................... 66
3.18 Alignment comparison with ICP. .................................................. 67
3.19 Shape profile of the 2nd layer. ................................................................. 67
3.20 3-level multi-resolution alignment results. ........................................... 68
3.21 Deviation profiles. ................................................................................. 68
3.22 Result of error quantification based on 100 uniformly sampled points. .... 69
3.23 Result of error quantification based on 57 points around the defective area. 70
3.24 Feature segmentation by spectral clustering. ......................................... 71
3.25 Shape profile of the outer contour shape. .............................................. 71
3.26 3-level multi-resolution alignment results based on the datum feature. .... 72
3.27 Sample layers from the upper half of the human heart ......................... 73
3.28 Systematic error from layer 75 to layer 175 ......................................... 73
3.29 The ADD under different mean of post-change model ....................... 74
4.1 The overview of three different layers .................................................... 86
4.2 Region division in the NS method (boundaries represented by dashed lines). 87
4.3 Region division in the HSS method (perimeter sub-paths omitted for clarity). (a) Sub-path #11 and #9 are two raster sub-paths of case 1. Both of them span two regions. Sub-path #11 is assigned to the left region, and sub-path #9 is assigned to the middle region. (b) Sub-path #8 is a raster sub-paths of case 2, which spans two regions and is assigned to the left region. (c) No sub-paths span multiple regions. 88
4.4 LPT algorithm flow chart ................................................................. 92
4.5 MOA algorithm flow chart ............................................................ 94
4.6 Computing overlapped area for two sub-paths .................................. 95
4.7 Example of finding the minimum delay .............................................. 95
4.8 False positive example for collision detection. The shadowed area is the true area swept by the extruders. ................................................. 96
4.9 Example of identifying inspection time period ................................... 97
4.10 Example of identifying inspection time period ................................. 99
4.11 Finding the bounding rectangle to form the extended sweeping area .... 99
4.12 Sub-paths of the layer (where axis values indicate the position of layer on the workspace) ................................................................. 101

4.13 Graphic illustration of the scheduling result from the LPT heuristic (a) The toolpath path is shown in time (vertical axis) indicating no intersection between extruder trajectories. (b) The relative size of the extruder’s bounding rectangle to the layer is provided. ................................................................. 102

4.14 Result of LPT, MOA, NS and HSS for 3-extruder scheduling. Negative signs indicates that the path was reversed to avoid collision. Black indicates delay introduced to avoid collision. ................................................................. 103

4.15 The layer overview of Case 2. ................................................................. 105

4.16 Performance Comparison among LPT, MOA, NS, HSS for Case 2 .................. 105

4.17 The layer overview of Case 3 ................................................................. 106

4.18 Performance Comparison among LPT, MOA, NS and HSS for Case 3 ............... 107
## List of Tables

2.1 The performance of three rigid alignment methods. .............................................. 7
2.2 LOS inspection without datum for a part with \( n \) features ............................... 16
2.3 Threshold of shape inspection for test object 2a and 2b. ..................................... 24
2.4 Threshold of datum-based LOS inspection for test objects 2a and 2c. ................. 25
2.5 Threshold of non-datum-based LOS inspection for test objects 2a and 2c. .......... 26
2.6 Threshold of LOS inspection for test object 3. ..................................................... 31
2.7 LOS inspection result for the test object 3 indicating the pose problem with feature
   #4 relative to all other features. ............................................................................ 31
2.8 Threshold of LOS inspection for test object 4a, 4b and 4c ................................. 32
3.1 Point-wise Deviation Based on the Multi-resolution Alignment .......................... 56
3.2 Result of ANOVA ................................................................................................. 69
3.3 Error Quantification of Each Feature ..................................................................... 71
4.1 Symbols Used in the Optimization Model .............................................................. 90
4.2 Performance of LPT and MOA for Case 1 vs Number of Extruders) .................. 101
4.3 Performance Comparison among LPT, MOA, NS and HSS for the Simple Layer .. 102
List of Published Papers


1 Introduction

Additive manufacturing (AM) is a form of creating a three-dimensional object layer by layer directly from a 3D CAD model. It now has been involved in many different industrial sectors to fabricate objects that have arbitrarily complex geometric features and can hardly be accomplished by traditional manufacturing processes. AM also shows its advantages in reducing material waste, improving part consolidation, and producing parts directly without the need for expensive part-specific tooling. However, many technical challenges have to be solved before AM reaches its full potential. One noticeable issue is the limited geometric accuracy of AM fabricated parts, which is still not sufficient to meet the requirements of some industries, such as biomedical or aerospace engineering. However, the traditional geometric inspection techniques might not perform well and even cannot be adopted in AM due to the geometric complexity of parts and additive layer-by-layer nature of AM processes. Moreover, the AM processes are no longer limited in producing small but complex geometries due to the growing demand for the low-volume fabrication of large products, especially in aerospace, architecture, and automotive industries. Nevertheless, upscaling the AM process is hindered by limited fabrication speed. One straightforward solution is to increase layer thickness, but this action sacrifices resolution and causes poor surface finish and dimensional accuracy. Aiming to address these limitations of AM processes, three research problems are illustrated in Figure 1.1 which are discussed and resolved in this work.

First, for offline inspection, the coordinate measuring machine (CMM) using the geometric dimensioning and tolerancing (GD&T) paradigm, which is applied in many traditional production environments, does have drawbacks in the context of AM. CMM inspection relies on datums established in the part design; however, this can be far from straightforward for AM designs, which in most cases lack parametric datum features. In addition, current AM techniques have shown the capability of producing complex or large-scaled AM parts with multiple fabrication cycles, or producing working assemblies in a single build, or printing on a pre-made surface. Each subpart, or so-called feature, is considered as an individual object of interest. In practice, the functionality of an AM produced part is dependent on the quality of features as well as their relative positioning.
Therefore, even if convenient datum surfaces exist in AM parts, designers are often interested in the relative position and orientation between two freeform features, rather than their individual poses with respect to a datum frame. Another drawback for CMM and GD&T paradigm is lacking accessibility for users outside of factory or laboratory environments. Fortunately, 3D scanning, as an alternative of CMM, can provide fast, high-density measurements and has now been widely applied in AM processes for geometric inspection. Based on the point cloud data obtained from 3D scanning, this research proposes a new feature-based scale- and pose- invariant quality inspection method using a novel Location-Orientation-Shape (LOS) distribution. The key technique of the new method is to describe the shape and pose of features via kernel density estimation and detect non-conformities based on statistical divergence. In section 2, technical details of this novel offline fast inspection method are provided and validated with numerical examples using both simulated and physical AM builds.

Second, due to the additive nature of AM, the online inspection should be blended with the offline geometric inspection to prevent errors of precedent layers getting accumulated on subsequent layers. The geometric accuracy of layers can be described by in-plane shape deviation. The shape of each inspected layer can be described by a 2D point cloud obtained by slicing a thin layer of 3D point cloud acquired from 3D scanning. In practice, a scanned shape must be aligned with the corresponding base-truth CAD model before evaluating its geometric accuracy. Then the shape deviation is described by point-wise deviation based on the aligned point clouds. Therefore, the calculated shape deviation is attributed to systematic, measurement, and alignment errors. However, how to distinguish errors from different sources has not been addressed in the previous studies. To fill this research gap, a multi-resolution point cloud alignment and inspection scheme based on wavelet decomposition and randomized completed block design (RCBD) is proposed in section 3. Based on the quantified alignment error, the performance of alignment is improved by a two-step optimization model. Moreover, since the systematic error is the one that triggers an alarm of system anomalies, a quickest change detection (QCD) algorithm is applied under the multi-resolution alignment and inspection framework to identify the layer where the earliest sys-
tematic deviation distribution changes during the printing process. Numerical experiments and a case study on a human heart is conducted to illustrate the performance of the proposed method in detecting the layer-wise geometric error.

Third, with demand for the low-volume fabrication of large products in industries such as aerospace, architecture, and automotive, there is a strong motivation to scale up the process. However, upscaling AM processes is restricted by the limited fabrication speed, especially for AM processes like Fuse Filament Fabrication (FFF). In practice, increasing layer thickness can speed up the process, but this sacrifices resolution and increases stair-stepping errors. This limitation forces manufacturers and customers to choose between long cycle times and extremely poor surface finishes and dimensional accuracy. Therefore, the fabrication efficiency of AM is desired to be improved without sacrificing the geometric quality. Fortunately, the idea of decomposing a 3D geometry into features provides a possibility of making multiple machines work collaboratively in parallel to accomplish a complex part. In this work, a general toolpath allocation and scheduling methodology is proposed in section 4 to accomplish the optimization of a concurrent AM process by distributing the fabrication tasks of each layer among multiple extruders. Two heuristic algorithms are proposed to solve the optimization model. The performance of the algorithms with respect to fabrication time and computational cost are demonstrated and compared using three case studies.

![Figure 1.1: An overview of three research problems](image)
2 Scale and Pose Invariant Feature-based Quality Inspection

Additive manufacturing (AM) has the unprecedented ability to create customized, complex, and nonparametric geometry, and it has made this ability accessible to individuals outside of traditional production environments. Geometric inspection technology, however, has yet to adapt to take full advantage of AM’s abilities. Coordinate measuring machines are accurate, but they are also slow, expensive to operate, and inaccessible to many AM users. On the other hand, 3D-scanners provide fast, high-density measurements, but there is a lack of feature-based analysis techniques for point cloud data. There exists a need for developing fast, feature-based geometric inspection techniques that can be implemented by users without specialized training in inspection according to geometric dimensioning and tolerancing conventions. This research proposes a new scale- and pose- invariant quality inspection method based on a novel Location-Orientation-Shape (LOS) distribution derived from point cloud data. The key technique of the new method is to describe the shape and pose of key features via kernel density estimation and detect non-conformities based on statistical divergence. Numerical examples are provided, and tests on physical AM builds are conducted to validate the method. The results show that the proposed inspection scheme is able to identify form, position, and orientation defects. The results also demonstrate how datum features can be incorporated into point cloud inspection; that datum features can be complex, nonparametric surfaces; and how the specification of datums can be more intuitive and meaningful, particularly for users without special training.

2.1 Introduction

Additive manufacturing (AM) has shown notable potential for agile production of complex freeform geometric features. Application examples include topology-optimized structural components (Brackett et al., 2011), prosthetics (Jin et al., 2015), and biological tissues and organs (Murphy and Atala, 2014). AM is highly agile, offering good potential for economical low-volume production and the ability to easily scale builds on demand. AM has also democratized design and production by allowing people without specialized engineering skills or access to capital equipment to design
and manufacture objects. Despite these advantages, geometric quality inspection techniques that support these strengths are required before AM can reach its full potential.

The gold standard for precision metrology in many production environments is a coordinate measuring machine (CMM) using the geometric dimensioning and tolerancing (GD&T) (Meadows, 2009) paradigm, but this does have drawbacks in the context of AM. CMM inspection tends to be highly manual, especially in high-mix, low-volume production where AM can be the most impactful. Aside from making inspection slow and expensive, this tends to result in relatively low sampling density focused on a few key features. But in a topology-optimized part such as that shown in Fig.2.1a, stresses are highly sensitive to geometric deviations throughout the part, necessitating a more complete inspection. CMM inspection also relies on datums established in the part design; however, this can be far from straightforward for designs which lack parametric features such as the one depicted in Fig.2.1b. Even if convenient datum surfaces exist, designers are often interested in the relative position and orientation between two freeform features, rather than their individual poses with respect to a datum frame, (e.g., the relative positions of teeth in a dental prosthesis (Juneja et al., 2018)). Moreover, CMM inspection for complex geometries often requires specialized tooling to fixture parts, undercutting the advantages of AM’s toolless nature, and the process is not scale invariant, making it more difficult to take advantage of AM’s ability to easily scale builds on demand. Finally, CMMs are relatively inaccessible outside of factory or laboratory environments, and establishing and inspecting against datums requires specialized knowledge, which are limiting factors for makers and entrepreneurs.

In contrast, 3D scanning is capable of high-density sampling of an entire part in a relatively short time without fixturing. It can also be cheaper and therefore more accessible outside of formal production environments, with many consumer-grade models under $1000 (Anonymity, 2017a). In practice, however, the scanned point cloud must be aligned with the corresponding base-truth CAD model prior to evaluating the geometric accuracy. Common alignment methods coming with a point cloud processing software package rely on least-square fitting algorithms, such as iterative closest point (ICP) algorithm (Besl and McKay, 1992), to find the optimal rigid transformation
that minimizes the sum of squared distances between the measured alignment points and their target reference points. However, it is possible that defects of the sample part might deteriorate the alignment accuracy, or cause the misalignment even around the non-defective area. This is illustrated in Fig.2.2, where a deviation in the ears of the Stanford bunny model (Turk and Levoy, 2005) results in misalignment that indicates geometric deviations in the head and foot features. To better justify this point, two more rigid registration methods based on Normal Distributions Transform (NDT) (Magnusson, 2009) and Coherent Point Drift (CPD) (Myronenko and Song, 2010) are implemented. The Root Mean Square Error (RMSE) of three different rigid alignment methods are provided in Table 2.1. It is observed that the influence of defects on the alignment result is significant. Indeed, there are many more advanced alignment methods in the literature, which are proposed for object recognition or 3D perception. For instance, local features-based alignment methods (Eckart et al., 2018) can resolve some shortcomings of the global ICP. However, the geometric accuracy of the selected features on a defective object are not able to be guaranteed. If the points associated with the selected feature have large deviations, then the alignment based on the feature apparently will still cause significant RMSE. Compared to rigid registration methods, non-rigid registration (Mateus et al., 2008) can perform a more complicated transformation that might change the shape of the point cloud to achieve better alignment. From the perspective of geometric inspection, the point cloud should not be changed or deformed in the registration
process. Therefore, in most literature or software associated with the shape inspection based on the point cloud data, only rigid registration methods are applied. (Hong-Seok and Mani, 2014; Li and Guskov, 2005; Shi et al., 2006; Minetola, 2012).

Figure 2.2: Example of misalignment using the ICP method.

Table 2.1: The performance of three rigid alignment methods.

<table>
<thead>
<tr>
<th></th>
<th>RMSE of a conforming sample</th>
<th>RMSE of a defective sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICP</td>
<td>0.69</td>
<td>2.38</td>
</tr>
<tr>
<td>NDT</td>
<td>1.02</td>
<td>2.26</td>
</tr>
<tr>
<td>CPD rigid</td>
<td>0.81</td>
<td>2.53</td>
</tr>
</tbody>
</table>

This research introduces a novel point cloud quality inspection method that blends many of the advantages of CMM/GD&T and 3D-scanning-based approaches while mitigating some of the disadvantages. The theoretical contribution is a generalization of the shape distribution concept used in artificial intelligence for object recognition. This generalization is dubbed the location-orientation-shape (LOS) distribution and can describe the relative form and pose between features represented by point clouds. Methods for form, location, and orientation inspection are developed based on the LOS distribution and tested via physical experiments on 3D printed parts. Since the LOS distribution is a piece-wise pdf and is no longer associated with a specific Cartesian Coordinate system (CCS), or we say a specific pose and scale of the part, alignment is not required before inspection to eliminate the effect of alignment error. It is a scale- and pose-invariant approach that can incorporate intuitively defined datums based on complex and/or nonparametric features. It is also able to yield meaningful, high-level inspection results even without datum features. Although
we propose an alternative to CMM/GD&T, we do not claim to replace or supersede it, especially for traditional parametric geometries. Rather, we propose a method that has different strengths and weaknesses which can supplement or be used in place of CMM/GD&T when warranted by the application. Practical applications of the proposed method include a first-order inspection for reducing the number of samples sent to CMM; an alternative inspection method for freeform geometry that is difficult to tolerance and/or establish datums; and an alternative for non-engineers who are not well versed in the intricacies of GD&T or have access to a CMM.

2.2 Literature Review

GD&T has limitations in tolerancing freeform geometries. Two major GD&T standards are ASME Y14.5 (Meadows, 2009) and ISO 1101 (Standard, 2004). In ASME Y14.5, manufacturing variations are categorized into different classes with specific tolerance types (e.g., form, orientation, location, and profile). Among all defined tolerance types, the profile tolerance is the one which has potential for capturing the dimensional accuracy of freeform surfaces. The profile of a surface describes a 3D tolerance zone around a surface by two offset surfaces. The size of the tolerance zone could be either uniform or non-uniform along the surface. The non-uniform tolerance zone can potentially be useful by adopting variable tolerances along the surface (Ameta et al., 2015b). However, it is not clear in the standard how to specify such a tolerance zone and its parameters (Savio et al., 2007). Moreover, an AM-produced freeform part cannot simply be modeled as a continuous surface by a mathematical formulation (Yamaguchi, 2012). In practice, a freeform part can be decomposed into multiple surfaces for the ease of formulation, but the position information of these surface segments essential to the functionality of the part, is unfortunately neglected. In addition, there are some other AM-relevant issues associated with GD&T (Ameta et al., 2015a) including (1) the feature and tolerance information lost during tessellation, (2) the lack of analysis in terms of building assemblies as a single component, (3) the error accumulations among layers, (4) the influence of datum reference in each layer, and (5) lack of methods for processing measurements obtained by non-CMM techniques. The method proposed in this work does not require
datum references and is able to process the point cloud measurement to quickly evaluate the quality of the part.

Motivated by the limitation of GD&T in tolerancing freeform geometries, alternatives have been proposed based on the result of non-contact measuring techniques (e.g., 3D scanner). Prior work in the area of point-cloud-based inspection may be classified into two categories: (1) modeling the as-printed surface via non-parametric approaches and comparing statistics to corresponding in-control values, and (2) comparing the as-printed geometry to a base-truth CAD model and calculating point-wise, layer-wise, or surface-wise deviation directly. The Gaussian Process (GP) and its extended methods are the main statistical models adopted in the first category of literature (Del Castillo et al., 2015; Colosimo et al., 2014, 2015; Jahani et al., 2018). The main idea is to build a parametric model for a surface using a Gaussian process and then establish a control chart on the GP parameters. This method has proven effective for inspecting simple geometries such as cylinders’ and tunnels’ inner surface (Zhang et al., 2017). However, a freeform AM part has more complicated geometric features and surfaces. Additionally, such methods are not computationally efficient for large datasets like point clouds. The methods in the second category focus more on deviation modeling. In Choi et al. (2002), Silva et al. (2008), and El-Katatny et al. (2010) geometric error is quantified by measuring the position of specific landmarks on the part. For example, in (El-Katatny et al., 2010), the linear distances of 11 landmarks on a printed skull are measured and then compared to the values obtained from the nominal model. However, error estimation based on relatively few landmarks is unlikely to be as accurate as those quantified by calculating point-wise distances between two point clouds (Ciocca et al., 2009; Germani et al., 2010; Meakin et al., 2004), calculating the deviation from the printed surface to a nominal one using a polar coordinate system (Zhu et al., 2018), or using some commercial software to generate a deviation profile or color-map (Rao et al., 2016). One limitation of these works is that the fabricated geometry must be aligned to the nominal model, which implies that both must be of the same scale. To improve inspection efficiency and inspect scaled parts within a single inspection framework, scale and orientation invariant methods for error estimation are desirable. In addition, these works focus on quantifying
the dimensional accuracy of individual features and do not consider quality inspection in terms of the relative positioning of complex features. The inspection methods based on shape distribution proposed in this work are intended to address these issues.

The methods proposed in this research are based on the concept of a shape distribution, which is a shape descriptor originally proposed by Osada et al. (2002) for object recognition. A shape descriptor is a simplified representation of a 2D or 3D shape in the form of a mathematical formulation. Shape distribution is a probability distribution sampled from a shape function measuring geometric properties of the 3D model such as distance, angle, area, or volume between random surface points. It evaluates the similarity of the objects by comparing probability distributions. The method is implemented in three steps: (1) select a shape function to create a shape vector, (2) construct the shape distribution for the object based on the shape vector, and (3) compare the shape distribution to those of known objects. In the first step of this method, common shape functions include the D2, D3, A3 and D4, where D2 indicates the Euclidean distance between any two points in the point cloud, D3 measures the square root of the area of the triangle between three random points on the surface, A3 measures the angle between three random points on the surface of a 3D model, and D4 measures the cube root of the volume of the tetrahedron between four random points on the surface. Then the shape distribution is constructed by a piecewise linear function based on the histogram of the shape vectors. The distance between two distributions can be compared via many standard approaches. Examples include the Minkowski $L_N$ norms, Kolmogorov-Smirnov distance, Kullback-Leibler divergence, match distances, earth mover’s distance, and Bhattacharyya distance. Some comprehensive surveys (Kazmi et al., 2013; Tangelder and Veltkamp, 2004) have been conducted for different types of 3D shape descriptors and their computational complexity. Compared to other shape descriptors, shape distribution is a more robust, efficient, and transformation-invariant method to quantify the dissimilarity between 3D shapes. However, the original shape distribution describes only the property of a single geometry/feature, and it does not contain location or orientation information. The proposed methods inherit the advantages of shape distribution while adding the capability of encoding the relative
Defining features and segregating the point cloud is prerequisite to any feature-based inspection method. First, features should be defined based on the function of the part. For instance, each of the separable working assemblies, which are produced in a single build (Ameta et al., 2015a) or by using multi-stage AM processes (Monika Mahto, 2017), can be identified as a feature. In addition, 6-axis printing (Song et al., 2015) started to gain interest for its potential ability in improving quality and mechanical properties of the printed part. The possibility and performance of printing on a pre-made surface have been investigated (Pierson and Chivukula, 2018). In this case, the geometry printed in each fabrication cycle is recognized as a feature. Once features are defined, the point cloud must be segmented accordingly, either manually or by machine learning techniques such as Clustering (Zhang et al., 2008; Lu et al., 2016) and Deep Learning (Charles et al., 2017; Ben-Shabat et al., 2017). Different machine learning techniques for point cloud segmentation are reviewed by Grilli et al. (2017). In this work, to focus on the key idea of quality inspection, features are pre-determined by users. Then k-means clustering is applied to segment the point cloud into desired pieces. The clustering result is verified by the user to remove any inconsistencies in the cluster labeling. In the future, more intelligent methods can be embedded into the system without influencing the remainder of the underlying inspection method.

2.3 Theory

The novel location-orientation-shape (LOS) distribution extends the concept of a shape distribution to include the relative pose between features. Whereas a shape distribution is obtained based on points within a single point cloud, the LOS distribution is based on analogous metrics between two point clouds. Additionally, the LOS distribution considers all points within the subject point clouds, rather than random points as in Osada et al. (2002). Given point cloud $\mathcal{P}_i$, its LOS vector associated with a reference point cloud $\mathcal{P}_j$ is denoted as $LOS\cdot fun\{\mathcal{P}_i, \mathcal{P}_j\}$, where $fun$ is a geometric function selected by the user. Generalizing the D2, D3, A3 and D4 shape vectors, the
where $LOS-D2$ is a distance vector measuring distances between any two points, $LOS-D3$ measures the square root area of triangles formed by three points, $LOS-A3$ measures angles between two vectors formed by three points, and $LOS-D4$ measures the cube root of the volumes of a tetrahedron formed by four points. Given these definitions, the shape vector created in Osada et al. (2002) is a special case of the LOS vector when the target point cloud $P_i$ itself is considered as the reference point cloud. For instance, the D2 shape vector for point cloud $i$ is

$$LOS-D2(P_i, P_j) = \{d(p, q) \mid d(p, q) = (p - q)(p - q)^T, \forall p \in P_i, q \in P_j\}$$

Equations for the other shape functions follow similarly.

The LOS distribution is constructed from an LOS vector in three steps: 1) The LOS vector is standardized such that its components are in the interval $[0,1]$ to eliminate the effect of scale. 2) A histogram is constructed for the values in the LOS vector. 3) A piece-wise kernel estimate is obtained from the histogram, which can be determined directly using a kernel fitting algorithm for the selected kernel function. In this work, a standard Gaussian kernel function is chosen and applied in Matlab to perform the kernel density estimation. Given the LOS vector $LOS-fun(P_i, P_j)$, the
corresponding LOS distribution is denoted as $g_{ij} = g\{P_i, P_j\}$.

The distance between an LOS distribution generated from 3D scan information and the one generated from the base-truth model can be used as a quality metric. A smaller distance value indicates more similarity between the two distributions. The distance between any two LOS distributions can be measured by a number of methods: the $\chi^2$ statistic, Bhattacharyya distance, Minkowski $L_N$ norms, Kolmogorov-Smirnov (K-S) distance, and so forth (Puzicha et al., 1999). In particular, the $\chi^2$ statistic and Bhattacharyya distance are commonly used for comparing binned data (histograms). On the other hand, the Minkowski $L_N$ norms measures the average difference between two pdfs while K-S distance is preferred in measuring the maximum difference between two cumulative distribution functions (cdfs). In practice, users have the flexibility to choose any similarity measure. K-S distance is selected for this work without loss of generality. Given two LOS distributions $g_{ij}^{(1)}$ and $g_{ij}^{(2)}$ constructed from the point clouds of two sample parts, the K-S distance between the two distributions is calculated by

$$d_{ij} = \max_{x_k} |G_{ij}^{(1)}(x_k) - G_{ij}^{(2)}(x_k)|, \forall x_k \in [0, 1]$$

(2.6)

where $G_{ij}^{(1)}$ and $G_{ij}^{(2)}$ are cdfs corresponding to $g_{ij}^{(1)}$ and $g_{ij}^{(2)}$.

### 2.4 Methods

#### 2.4.1 Inspection based on LOS Distribution

A summary of the LOS quality inspection method is as follows (Fig.2.3): 1) A point cloud for the fabricated part is obtained via 3D scanning and segmented into features. Features may be parametric or non-parametric and should include datum features when defined. Each feature point cloud is downsampled using a Voxel Grid Filter (Anonymity, 2017b) to reduce the size of the data set and computation time. The result is a sampled feature point cloud $P_i^{(s)}$, for each feature $i$. 2) The base-truth CAD model is sampled to obtain a nominal point cloud, which is similarly segmented and downsampled. The result is a set of nominal feature point clouds $P_i^{(n)}$. 3) LOS
distributions $g^{(s)}$ and $g^{(n)}$ are computed based on these point clouds. The specific point cloud combinations used to form $g^{(s)}$ and $g^{(n)}$ depend on the inspection objective, as detailed below. The geometric quality metric for a given feature is quantified by the distance, $d$, between $g^{(s)}$ and $g^{(n)}$.

**LOS Shape Inspection.** When an LOS vector is constructed for a point cloud with respect to itself, as with Eq.5, the resulting distribution contains information about that point cloud’s shape. Let $n$ denote the number of features to be inspected and $i$ be a feature index where $i \in \{1, 2, ..., n\}$. The distance, $d_{ii}$, between LOS distributions $g^{(s)}_{ii} = g\{\mathcal{P}^{(s)}_i, \mathcal{P}^{(s)}_i\}$ and $g^{(n)}_{ii} = g\{\mathcal{P}^{(n)}_i, \mathcal{P}^{(n)}_i\}$ indicates shape deviation. This is analogous to form inspection in GD&T, but is not restricted to simple parametric geometries such as planes and cylinders.

**LOS Inspection with Datum Features.** When an LOS vector is constructed between a given point cloud and a reference point cloud, as in Eq.1-4, the resulting distribution contains information about the given point cloud’s location, orientation, and shape with respect to the reference. Let $n$ denote the number of inspection features, $m$ the number of datum features, $i$ the inspection feature index, and $j$ the datum feature index, where $i \in \{1, 2, ..., n\}$ and $j \in \{1, 2, ..., m\}$.

**Figure 2.3:** LOS inspection process overview.
The distance, $d_{ij}$, between LOS distributions $g_{ij}^{(s)} = g\{P_i^{(s)}, P_j^{(s)}\}$ and $g_{ij}^{(n)} = g\{P_i^{(n)}, P_j^{(n)}\}$ is a combined measure of the shape and pose of inspection feature $i$ with respect to datum feature $j$. Although datum features may be simple surfaces as is conventional in GD&T, inspection via LOS distributions allows the concept of a datum feature to be expanded to include complex surfaces and/or non-parametric surfaces.

The coupling between shape and pose information in LOS inspection can be advantageous in a go/no-go inspection context because it requires only one set of computations and yields a single numeric result. In other contexts, however, it may be desirable to have separate information about shape and pose. In this situation, shape inspection may be combined with LOS inspection. If a feature’s shape deviation, $d_{ii}$, is low, but its LOS deviation, $d_{ij}$, is high, one may conclude that the primary problem with the feature is its location and/or orientation.

**LOS Inspection without Datum Features.** Outside of industrial production environments, models intended for AM often lack defined datum features, yet there is still a need for meaningful, quantitative geometric quality assessment. In this case, LOS inspection may be applied between all features of interest. Let $n$ denote the number of features of interest and $i, j$ are the indices for a pair of features where $i \in \{1, 2, ..., n\}$ and $j \in \{1, 2, ..., n\}$. The distances, $d_{ij}$, between LOS distributions $g_{ij}^{(s)} = g\{P_i^{(s)}, P_j^{(s)}\}$ and $g_{ij}^{(n)} = g\{P_i^{(n)}, P_j^{(n)}\}$ form an LOS inspection matrix, as shown in Table 2.2.

The diagonal contains the shape inspection results for all features. Each row corresponds to one feature, with the elements representing the LOS inspection results with respect to each of the other part features. Any large element in the matrix is cause for rejecting the part, but more importantly, the LOS inspection matrix gives the user insight into why the part is rejected. For instance, if $d_{33}$ and $d_{67}$ are two elements of the matrix that give cause to reject the part, the user may deduce that 1) feature #3 is not well shaped and 2) feature #6 is not positioned accurately with regard to feature #7. This information can prove valuable in diagnosing the cause of the failure and devising a means to correct the problem.

It is worth noting that the threshold on distance values to differentiate acceptable and unaccept-
Table 2.2: LOS inspection without datum for a part with \( n \) features

<table>
<thead>
<tr>
<th></th>
<th>( P_1 )</th>
<th>( P_2 )</th>
<th>...</th>
<th>( P_l )</th>
<th>...</th>
<th>( P_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>( d_{11} )</td>
<td>( d_{12} )</td>
<td>...</td>
<td>( d_{1i} )</td>
<td>...</td>
<td>( d_{1n} )</td>
</tr>
<tr>
<td>( P_2 )</td>
<td>( d_{21} )</td>
<td>( d_{22} )</td>
<td>...</td>
<td>( d_{2i} )</td>
<td>...</td>
<td>( d_{2n} )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>( P_l )</td>
<td>( d_{l1} )</td>
<td>( d_{l2} )</td>
<td>...</td>
<td>( d_{2i} )</td>
<td>...</td>
<td>( d_{ln} )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>( P_n )</td>
<td>( d_{n1} )</td>
<td>( d_{n2} )</td>
<td>...</td>
<td>( d_{ni} )</td>
<td>...</td>
<td>( d_{nn} )</td>
</tr>
</tbody>
</table>

Viable parts can be determined by the distribution of divergence considering natural variation of AM process and user-defined tolerance. Technically, the threshold can be obtained based on simulated models which are created by adding only random deviations to the nominal model. The random deviation follows a normal distribution which has a zero mean and a standard deviation equal to the worst variability a user can tolerate and determine. After the distributions of K-S distances based on simulated models are established, the threshold of a certain K-S distance value, in this work, is defined by \( \mu_d + 3\sigma_d \), where \( \mu_d \) and \( \sigma_d \) are denoted as the mean and standard deviation of the K-S distance distribution obtained from simulated models. Specifically, if the observed deviation (i.e., K-S distance) is larger than the threshold, then the part is considered as an unacceptable sample.

Feature indexing. To this point it has been tacitly assumed that feature indexing is consistent between the sampled and nominal point clouds, \( P^{(s)} \) and \( P^{(n)} \). If feature definition is a manual process, with or without assistance from the automated feature segmentation algorithms mentioned in section 2, this is a reasonable assumption; however, inconsistent indexing could be problematic in fully automated feature definition and would result in false rejections. In such circumstances, comparison of the LOS distributions can assist in establishing the correct feature correspondence.

The procedure begins by matching and re-indexing similarly shaped features: Given feature \( j \)
in the nominal model, the corresponding feature in the scanned sample can be determined by

\[ i = \arg \min_i (\max |G_{ii}^{(s)} - G_{jj}^{(n)}|) \tag{2.7} \]

where \( \max |G_{ii}^{(s)} - G_{jj}^{(n)}| \) measures the K-S distance between the shape-based LOS distribution of the given nominal feature \( j \) and any sample feature \( i \). For any feature \( j \), if the corresponding \( i \neq j \), then feature \( i \) in the scanned model is re-indexed as \( j \). The matched results can further be interpreted in the context of LOS shape inspection as discussed above. If the shape inspection results are acceptable, and if all features are uniquely shaped, then the indexing result is unambiguous and LOS inspection can proceed; however, substantially similar features may still be mis-indexed.

LOS distributions may further be used to resolve the indexing ambiguity caused by features with similar form if one correctly indexed feature can be identified. This can be done automatically if at least one feature has a unique shape. The unique features can be differentiated based on shape distribution comparisons. The procedure proposed in Osada et al. (2002) can be implemented by considering one feature as the test shape and the rest as the reference shape database. If the shape distribution of the test shape cannot be grouped into the shape distribution of any reference shape, then the test shape is considered as a unique feature. If no unique features exist, user intervention (via an appropriately designed interface) is required to identify one correctly indexed feature.

Once a correctly indexed feature has been identified, let \( i \) be that feature’s index. Given feature \( l \) of the nominal model, the corresponding feature in the scanned point cloud can be determined by

\[ k = \arg \min_k (\max |G_{ik}^{(s)} - G_{il}^{(n)}|) \tag{2.8} \]

where \( \max |G_{ik}^{(s)} - G_{il}^{(n)}| \) is the K-S distance between two pose-based LOS distributions, measuring the difference between the position of sample feature \( i \) relative to feature \( k \) of the scanned sample and position of nominal feature \( i \) with regard to feature \( l \) in the nominal model. If \( k \neq l \), then feature \( k \) in the scanned sample is re-indexed as \( l \). The rationale: If feature \( k \) in the scanned sample is actually feature \( l \) of the nominal model, their position relative to a correctly indexed feature \( i \) should be similar.
2.4.2 Point Cloud Downsampling

Due to the large size of the raw point cloud generated by a 3D scanner, downsampling is desirable to reduce the size of the dataset and improve computational efficiency (Steinberg, 2016). A Voxel Grid Filter (VGF) is used as a uniform sampling strategy that superimposes a voxel grid of user-defined resolution over the input point cloud and approximates all points within each voxel using their centroid.

In the original application of the shape distribution (Osada et al., 2002), it was shown that variation in point cloud density does not affect the overall curvature of the distribution, but it can generate small local differences which are undesirable for quality inspection. Even though a dense point cloud can provide more details on the distribution curve, the computation cost forces users to make a trade-off between the computational efficiency and inspection accuracy. In practice, for a point cloud with thousands of points, the LOS distribution can be constructed in seconds.

More importantly, since the proposed method is designed for scale-invariant quality inspection, the nominal model and the fabricated sample do not need to have the same scale. In this case, the voxel grid sizes chosen to downsample the nominal and fabricated point clouds should result in similar point cloud densities. Mathematically, a point cloud is represented as an $N \times 3$ matrix, where $N$ is the number of points in the point cloud. Let $\mathcal{OP}_i^{(s)}$ and $\mathcal{OP}_i^{(n)}$ denote the original point cloud of feature $i$ generated from the 3D scanner and the nominal model, respectively. Then the corresponding downsampled point clouds are denoted as $\mathcal{DP}_i^{(s)}$ and $\mathcal{DP}_i^{(n)}$. Assume the downsampled nominal point cloud $\mathcal{DP}_i^{(n)}$ has density $N_i^{(n)}$, then the downsampling grid size $\delta$ for the scanned point cloud $\mathcal{OP}_i^{(s)}$ should satisfy

$$\delta^* = \arg \min_{\delta} |N_i^{(n)} - N_i^{(s)}|,$$

where $N_i^{(s)}$ is the density of downsampled scanned point cloud of feature $i$ using VGF under grid size $\delta$. Since the two downsampled point clouds might not have equal density, the optimal grid size $\delta^*$ is thus selected for achieving the minimum difference between the densities of the two
downsampled point clouds.

2.4.3 Experimental Description

Four experiments were conducted to evaluate the proposed methods on both parametric and non-parametric geometries. Each considers one of the test objects in Fig.2.4.

Experiment 1 is designed to illustrate the sensitivity of the LOS distribution to pose, scale, and point cloud density. Based on the prior knowledge, the LOS distribution of a point cloud with respect to itself is expected to be scale and pose invariant and robust to point cloud density in quality inspection. A nominal point cloud is generated from the digital model for test object 1. This is then manipulated to simulate scans of rotated and scaled prints, and the resulting LOS distributions are compared to nominal distribution. LOS distributions of point clouds of dissimilar density are also compared.

Experiment 2 is designed to test the LOS inspection method on an object with parametric features. A geometrically correct copy of test object 2 is printed along with variations that serve as examples of errors. LOS shape inspection, inspection with a datum feature, and inspection without a datum feature are performed on each. Indexing disambiguation is also tested for four identical features.

Experiment 3 is designed to further implement the LOS inspection and feature re-indexing on a topology optimized object (Anonymity, 2019) as shown in Fig.2.4(c). A simulated defective sample is tested by following the procedure of LOS inspection without datum features as in Experiment
Experiment 4 is designed to test the LOS inspection method on an object with non-parametric features. Variations of test object 3 are printed to represent good and bad parts, and LOS inspection is performed without a datum feature. Moreover, the proposed methods are applied to biomedical AM application in Experiment 4 to illustrate its practical value and potential future application. In recent years, the demand of artificial organs or arthrosis continues to increase due to the lack of suitable donors. Different from objects for industrial purposes, biomedical parts usually consist of small components like vessels, scaffolds, bones and so on. The position accuracy of each component might impact the functionality of the final product. Therefore, a knee arthrosis shown in Fig.2.4(d) is designed as the third test object to inspect the shape and relative pose of two knee bones.

The LOS-D2 distribution was selected for all experiments to minimize computational expense. Physical artifacts for Experiments 2 and Experiment 4 were printed on a MakerGear FFF machine with 1.75mm PLA filament and scanned using a HDI 100 series 3D scanner from LMI Technologies. It is important to note that in Experiments 2 and Experiment 4, no attempt was made to fixture or align the physical artifact or the resulting point cloud with respect to the nominal model.

2.5 Results

2.5.1 Experiment 1

The first experiment is designed to demonstrate the sensitivity of LOS distribution for a point cloud with respect to itself in terms of pose, scale and point cloud density, and to verify the advantage of LOS distribution in quality inspection. In this experiment, the entire test object is considered as a single feature. There are three sample point clouds in Fig.2.5(b)-(d) generated by simulation based on the given nominal point cloud which is downsampled with grid size equal to 2 as shown in Fig.2.5 (a). All three test objects 1a-1c are designed to simulate the situation in reality such that the scanned point cloud obtained from the 3D scanning 1) has different point cloud density from the nominal model, 2) has different pose from the nominal model, or 3) has different scale from the
nominal model. The LOS distribution comparison between each sample and the nominal model are shown in Fig.2.6. The threshold to claim an unacceptable sample is equal to 1.9796, which is established based on 20 simulated models with a worst variability equal to 0.5 unit.

Figure 2.5: The point cloud associated with test object 1 was modified for scale, orientation, and point cloud density to test the sensitivity of the LOS distribution to these factors.

The first sample point cloud shown in Fig.2.5(b) is downsampled with grid size equal to 6, and thus it is less dense than the nominal point cloud. The K-S distance between their corresponding LOS distributions is 0.572. The K-S distance of the first sample is smaller than the threshold but larger than the other samples. This illustrates the importance of maintaining similar point cloud density between the nominal and scanned point clouds to reduce the variation observed in the K-S distance. In practice, point cloud density is determined by various inspection objectives (Colosimo et al., 2010; Huang et al., 2018). In some cases, sufficient density is required to capture high-spatial-frequency information of the point cloud; in other cases, low resolution is preferred considering the computational efficiency. The LOS distribution can be applied in both cases as long as Eq.(9) is satisfied to ensure that same level of details is maintained in sample and reference.
Figure 2.6: LOS distributions comparison for test object 1 are invariant to pose and scale, but are sensitive to point cloud density.

The second sample shown in Fig.2.5 (c) has the same scale and downsampling grid size as the nominal model, but its orientation is altered by rotating the object by 180, 160 and 90 degrees along the x,y, and z axes, respectively. The K-S distance between the LOS distribution of the second sample point cloud and nominal point cloud is 0.092, indicating that the LOS distribution is not sensitive to the pose of the part.

In Fig. 2.5 (d), the third sample part is scaled down to 30% of the nominal size and its grid size is set according to Eq. 2.9. The K-S distance between the LOS distribution and the nominal point cloud is 0.101. This indicates that the LOS distribution is scale-invariant as long as similar point cloud densities are maintained between parts.

The results of the second and the third sample indicate that the LOS distribution, as a shape profile, is invariant to the pose and scale of the shape. Each point cloud will have a unique LOS distribution, which is a piece-wise pdf and is no longer associated with a specific CCS, or a specific pose and scale of the part. Therefore, it is claimed that by the proposed inspection method based
on the LOS distribution, it does not require alignment.

2.5.2 Experiment 2

The second experiment is conducted on three test objects shown in Fig.2.7(a)-(c). Each test object contains four features of interest which are indicated in Fig.2.7(d). The point clouds of these features are obtained by k-means clustering and are indexed as shown in Fig.2.7(d). Test object 2a is printed from an unmodified CAD file. The test objects 2b and 2c were printed from modified CAD files to create shape and pose deviations that might occur during the printing process, as shown in Fig.2.7(b) and 2.7(c). Point clouds are downsampled with grid size equal to 1.8. The thresholds for each inspection process are calculated based on 20 simulation models with a worst variability equal to 0.1 unit.

![Figure 2.7: Objects printed for experiment 2(a-c); features of interest for experiment 2.](image)

**LOS Shape Inspection of Features.** To test the effectiveness of the proposed method in shape inspection, the LOS distribution comparison is conducted on test objects 2a and 2b. This tests the methods’ ability to detect the feature shape error present in test object 2b. The thresholds for shape inspection of each feature are presented in Table 2.3. The shape deviation of two test objects are shown in Fig.2.8. The results for test object 2a show that the LOS distributions generated from the sample part nearly overlap those based on the nominal model. Correspondingly, K-S distances of all four features are smaller than the thresholds indicating that all of them are acceptable. In contrast, in the results for test object 2b show that the shape deviation of feature #3 is signifi-

23
cantly larger than the threshold, indicating that this feature is not acceptable. It is worth noting that
the potential application of LOS shape inspection in detecting non-homogeneous shape shrink-
age/expansion is indicated in the result of test object 2b, which has a defective feature (feature #3)
with abnormal shape expansion.

Table 2.3: Threshold of shape inspection for test object 2a and 2b.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature 1</td>
<td>1.471</td>
</tr>
<tr>
<td>Feature 2</td>
<td>1.496</td>
</tr>
<tr>
<td>Feature 3</td>
<td>1.472</td>
</tr>
<tr>
<td>Feature 4</td>
<td>1.515</td>
</tr>
</tbody>
</table>

Figure 2.8: Shape inspection results for test objects 2a and 2b detect the shape defect introduced in test object 2b.

LOS Inspection with Datum Features. LOS inspection with datum features is applied to test objects 2a and 2c. This tests the method’s ability to detect the feature pose error in object 2c. The
LOS deviation for each feature of interest is measured relative to the datum feature indicated in
Fig.2.9. The thresholds for datum-based LOS inspection are presented in Table 2.4. The results of test objects 2a and 2c are shown in Fig.2.10. Sample and nominal LOS distributions nearly overlap in every case except for feature #3 of object 2c. The corresponding K-S distances confirm that the shape and/or pose of this feature is significantly larger than its corresponding threshold, and is
therefore not acceptable. To further narrow down the problem with this feature, a shape inspection is performed. Given that the shape inspection results for this feature are similar to those observed in Fig.2.8, one may conclude that pose is the primary issue with feature #3 of test object 2c.

![Figure 2.9: Datum feature and four target features.](image)

Table 2.4: Threshold of datum-based LOS inspection for test objects 2a and 2c.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature 1</td>
<td>2.224</td>
</tr>
<tr>
<td>Feature 2</td>
<td>2.399</td>
</tr>
<tr>
<td>Feature 3</td>
<td>2.092</td>
</tr>
<tr>
<td>Feature 4</td>
<td>2.269</td>
</tr>
</tbody>
</table>

**LOS Inspection Without Datum Features.** LOS inspection is also conducted on test objects 2a and 2c to test the method’s ability to detect the feature pose error in object 2c without referencing a datum feature. For each pair of features, feature $P_i$ and feature $P_j$ for $i, j \in \{1, 2, 3, 4\}$, the nominal and sample LOS distributions are plotted and the corresponding inspection matrices are shown in Fig.2.11 for object 2a and in Fig.2.12 for object 2c. The full matrices are shown for completeness, but analysis is focused on the upper triangle due to their symmetry. The threshold for each entry in the K-S distance table are presented in Table 2.5. For object 2a, all nominal and sampled LOS distributions appear to overlap, and none of the corresponding K-S distances stand out. The same is true of the diagonal values for object 2c, indicating that feature shapes are acceptable; however, $d_{13}$, $d_{23}$ and $d_{34}$ clearly stand out, indicating that pose errors exist in the part. For this test case, the inspection matrix yields insight into what is wrong with the part. The interpretation is that features 1 & 3, 2 & 3, and 3 & 4 are not correctly positioned relative to each other. Given that all features except for 3 appear to be correctly positioned relative to one another, it is deduced that feature 3 is incorrectly positioned.
Figure 2.10: LOS inspection result for test objects 2a and 2c detect the pose defect introduced in test object 2c.

Table 2.5: Threshold of non-datum-based LOS inspection for test objects 2a and 2c.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>1.470</td>
<td>1.809</td>
<td>2.234</td>
<td>1.838</td>
</tr>
<tr>
<td>$P_2$</td>
<td>1.809</td>
<td>1.614</td>
<td>1.965</td>
<td>2.517</td>
</tr>
<tr>
<td>$P_3$</td>
<td>2.234</td>
<td>1.965</td>
<td>1.426</td>
<td>2.438</td>
</tr>
<tr>
<td>$P_4$</td>
<td>1.838</td>
<td>2.517</td>
<td>2.438</td>
<td>1.432</td>
</tr>
</tbody>
</table>

The results do highlight two weaknesses of LOS inspection without datum features. First, if multiple features are incorrectly positioned, the interpretation of which features are out of place might not be clear. Large values in the matrix would indicate pose problems within the part even if the specifics are not evident, but inspection against a datum feature offers a more definitive interpretation. Second, its computational expense scales in $\frac{n^2+n}{2}$, where $n$ is the number of features. In contrast, inspection against a datum feature scales in $n$. For these reasons, inspection against a
Figure 2.11: The LOS distribution comparison for test object 2a indicates no features out of place with respect to each other.

Figure 2.12: The LOS distribution comparison for test object 2c indicates pose problems with feature 3 relative to features 1, 2, and 4.
Datum feature is preferred.

**Feature Re-indexing.** The ability to disambiguate indexing for similar features was tested on test objects 2a and 2c. Features in the sample point clouds were renumbered as shown in Fig. 2.13. In the absence of unique features, one correct feature index must be known a priori; in this case, feature #2 is specified as being correctly indexed.

![Nominal model](image1)

(a) Nominal model  
![Test object 2a](image2)  
(b) Test object 2a  
![Test object 2c](image3)  
(c) Test object 2c

Figure 2.13: An example when feature indexing of nominal model and sample parts is inconsistent.

![Feature re-indexing](image4)

Referring to Eq. 2.8, $i$ represents this feature, $l \in \{1, 2, 3, 4\}$ represents the indices of features in the nominal point cloud, and $k \in \{1, 2, 3, 4\}$ represents those in the scanned point cloud. The results in Figs. 2.14 and 2.15 show all combinations of K-S distances, $(\max |G^{(s)}_{ik} - G^{(n)}_{il}|)$, for the test parts. The correct figure indexes for the sampled parts are the values of $k$ associated with the column-wise minimum. Taking test object 2a in Figs. 2.14 as an example, feature #1 ($l = 1$) in the

<table>
<thead>
<tr>
<th>$i = 2$</th>
<th>$G^{(n)}_{il}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 1$</td>
<td>$l = 1$</td>
</tr>
<tr>
<td>$k = 2$</td>
<td>15.87</td>
</tr>
<tr>
<td>$k = 3$</td>
<td><strong>1.27</strong></td>
</tr>
<tr>
<td>$k = 4$</td>
<td>22.66</td>
</tr>
</tbody>
</table>

Figure 2.14: Feature re-indexing was successful for test object 2a.
nominal model corresponds to feature #3 in the scanned model, feature #3 \((l = 3)\) in the nominal model corresponds to feature #4 in the scanned model, and so forth. The resulting renumbering of the scanned features for both parts is graphically depicted in Figs.2.14 and 2.15. The results are correct, even in the presence of the feature pose present in test object 2c.

### 2.5.3 Experiment 3

The third experiment is conducted on a topology optimized object, which has seven features of interest which are created by k-means clustering. The feature indexing of the nominal model and the test object are shown in Fig.2.16 (a) and (b). The test object is created by simulation with one feature (feature #1 in the test object which is feature #4 in the nominal model) deviated from its original position. It is observed that the feature indexing is not consistent between the test object and the nominal model. Therefore, the features of the test object should be re-indexed before LOS inspection. It is worth noting that all features presented in Fig. 2.16 (c) have similar shape distributions, further demonstrating the effectiveness of the proposed methods in dealing with similar features. Point clouds are downsampled with grid size equal to 0.01. The thresholds
for each inspection process are calculated based on 20 simulation models with a worst variability equal to 0.0001 unit.

Figure 2.16: Test objects for Experiment 3; the shape distribution of each feature is compared in (c).

First, the index of the test object 3 should be adjusted based on the feature indexing of the nominal model. In this case, feature #2 is identified as the feature being correctly indexed by shape distribution comparisons referring to Eq. 2.7. Due to the similarity of shape distributions of the rest features, their re-indexing is performed by following the procedure as described in Experiment 2 for non-unique features. The re-indexed features of the test object 3 are shown in Fig.2.17 based on the similarity of each pair of LOS distributions. Next, the shape and position of each feature is examined by LOS inspection without datum features. The threshold for each entry in LOS inspection table is presented in Table 2.6. After the re-indexing, the pre-known defective feature is feature #4. Therefore, in the LOS inspection result shown in Table 2.7, large off-diagonal entries associated with feature #4, which are greater than the corresponding thresholds, are observed, indicating the existence of the pose problem of feature #4 relative to others.
Figure 2.17: Feature re-indexing was successful for test object 3.

Table 2.6: Threshold of LOS inspection for test object 3.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
<th>$P_6$</th>
<th>$P_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>0.79</td>
<td>0.58</td>
<td>0.24</td>
<td>0.33</td>
<td>0.10</td>
<td>0.13</td>
<td>0.08</td>
</tr>
<tr>
<td>$P_2$</td>
<td>0.58</td>
<td>0.38</td>
<td>0.68</td>
<td>0.57</td>
<td>0.74</td>
<td>0.91</td>
<td>0.76</td>
</tr>
<tr>
<td>$P_3$</td>
<td>0.24</td>
<td>0.68</td>
<td>0.11</td>
<td>0.31</td>
<td>0.13</td>
<td>0.19</td>
<td>0.10</td>
</tr>
<tr>
<td>$P_4$</td>
<td>0.33</td>
<td>0.57</td>
<td>0.31</td>
<td>0.44</td>
<td>0.24</td>
<td>0.31</td>
<td>0.53</td>
</tr>
<tr>
<td>$P_5$</td>
<td>0.10</td>
<td>0.74</td>
<td>0.13</td>
<td>0.24</td>
<td>0.60</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>$P_6$</td>
<td>0.13</td>
<td>0.91</td>
<td>0.19</td>
<td>0.31</td>
<td>0.09</td>
<td>0.37</td>
<td>0.35</td>
</tr>
<tr>
<td>$P_7$</td>
<td>0.08</td>
<td>0.76</td>
<td>0.10</td>
<td>0.53</td>
<td>0.06</td>
<td>0.35</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 2.7: LOS inspection result for the test object 3 indicating the pose problem with feature #4 relative to all other features.

<table>
<thead>
<tr>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
<th>$P_6$</th>
<th>$P_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.66</td>
<td>0.05</td>
<td>0.15</td>
<td>1.06</td>
<td>0.06</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>0.05</td>
<td>0.15</td>
<td>0.11</td>
<td>1.29</td>
<td>0.08</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>0.15</td>
<td>0.11</td>
<td>0.09</td>
<td>4.50</td>
<td>0.07</td>
<td>0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>1.06</td>
<td>1.29</td>
<td>4.50</td>
<td>0.41</td>
<td>5.77</td>
<td>1.14</td>
<td>1.08</td>
</tr>
<tr>
<td>0.06</td>
<td>0.08</td>
<td>0.07</td>
<td>5.77</td>
<td>0.58</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>0.07</td>
<td>0.11</td>
<td>0.14</td>
<td>1.14</td>
<td>0.08</td>
<td>0.35</td>
<td>0.31</td>
</tr>
<tr>
<td>0.07</td>
<td>0.01</td>
<td>0.04</td>
<td>1.08</td>
<td>0.04</td>
<td>0.31</td>
<td>0.13</td>
</tr>
</tbody>
</table>
2.5.4 Experiment 4

Two bones joined make up the arthrosis shown in Fig. 2.18. The lower half is considered as feature #1 while the upper half is feature #2. LOS inspection is conducted to inspect the shape and pose of these components. Test object 4a is printed with the intended geometry, while pose defects are deliberately introduced into test objects 4b and 4c. Point clouds and as-built artifacts are shown in Fig. 2.18 for all three parts. All test objects are compared to the nominal model which is downsampled with grid size equal to 3. In this experiment, the thresholds to differentiate acceptable/unacceptable test object are established based on 20 simulated models with a worst variability equal to 0.3 unit. The threshold for each of the four entries in the K-S distance table are presented in Table 2.8.

![Figure 2.18: Three test objects were printed for Experiment 4: one unaltered and two with pose errors introduced.](image)

<table>
<thead>
<tr>
<th>Threshold</th>
<th>$P_1$</th>
<th>$P_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>1.232</td>
<td>3.166</td>
</tr>
<tr>
<td>$P_2$</td>
<td>3.166</td>
<td>1.164</td>
</tr>
</tbody>
</table>

The LOS inspection result and K-S distance values of three test objects are presented in Fig. 2.19. It is observed that the test object 4a is accepted because all entries in the table are within their cor-
responding thresholds. Test object 4b is rejected due the large off-diagonal K-S distance value 4.21, which is greater than the corresponding threshold 3.166, indicating the existence of pose error of one feature relative to the other. Test object 4c is also unacceptable by following the similar rationale.

In addition, this experiment also gives insight that LOS distribution for a point cloud with respect to itself is pose-invariant. The diagonal entries in the table indicated the shape accuracy of a specific feature while the off-diagonal entries indicated the pose accuracy of a feature relative to another. From the diagonal entries associated with the shape of feature #2, it is observed that even the feature #2 of test objects 4b and 4c are altered, and the results of feature #2 shape inspection in the three tables are consistent in all three test objects. That is why we claim the shape inspection of LOS distribution is pose invariant.

2.6 Discussion

The results show that LOS distribution methods were able to differentiate between correctly and incorrectly printed features in physical trials. The results confirm that the method works for both parametric and nonparametric geometries, exhibits good pose and scale invariance, and that it can
resolve ambiguities caused by inconsistent feature indexing. It is worth noting that the process of establishing threshold based on simulation models will not increase the computational cost on a large scale. In most case, with reasonable sampling resolution, the threshold can be obtained within a minute. In addition, the simulation can be integrated with the inspection process in a “black box” for non-engineers. Therefore, the only input required from the users are scanned point cloud, the nominal model and the worst variability she/he can tolerate.

To evaluate the LOS distribution methods in the context of other more conventional methods, it is important to note that the former represents a fundamentally different, higher-level approach. Conventional methods focus on measuring the deviation of individual points from a CMM or 3D scanner. In the case of GD&T inspection, these points are typically classified as lying inside or outside of a volume that defines the tolerance for a given feature. In other schemes, the distance from each point to the ideal surface is quantified and analysis is based on this metric. In contrast, LOS inspection focuses on features as holistic entities represented by point clouds. It eliminates rigidly defined tolerance zones in favor of shape similarity measures. This alternative approach is not necessarily better or worse than conventional inspection methods; rather, its advantages and disadvantages must be considered with the specific inspection objectives in mind.

Key advantages of LOS inspection are related to datums and datum surfaces. Datums are conventionally established to serve as both a reference for measurement and to facilitate repeatable fixturing. As such, datum features are usually chosen to be simple parametric entities (planes, cylinders, etc.) and kept to the minimum required to meet 3-2-1 fixturing requirements. Although not strictly required, all critical part features usually reference a common set of datums to avoid refixturing during inspection. Such constraints are not a concern in LOS inspection because its insensitivity to the workpiece pose eliminate the need for fixture tooling. Each feature of interest can have unique datum(s), and since datums are represented via point clouds, they can be directly defined by complex combinations of parametric and/or nonparametric surfaces without resorting to implied datums. Time spent fixturing parts for inspection is eliminated, as is the cost and lead time associated with special fixturing. Features can also be inspected against datums that are of a
different scale. From the designer’s perspective, datum specification becomes more intuitive and potentially more meaningful. This brings feature-based inspection to AM users who are not trained in the intricacies of GD&T and work outside of formal manufacturing environments.

LOS inspection is based on 3D scanning, which yields high point cloud densities in considerably less time than CMM. It can also leverage x-ray CT to inspect internal structure and other areas inaccessible via CMM such as deep holes, undercuts, and so forth. Conventional analysis of 3D scanned objects begins by aligning the measured point cloud with the nominal model, which poses problems. First, this cannot be done when the measured point cloud is scaled differently from the nominal model. Second, point cloud alignment is usually accomplished by least squares or some other global optimization algorithm. Reference features are not necessarily aligned correctly with the corresponding points in the measured point cloud, so inspection relative to these features is problematic. LOS inspection solves both problems.

The scale invariance of the LOS inspection method has an important application in direct-write AM processes where post-print processing can introduce significant shrinkage. Examples include thermoset polymer matrix composites (Pierson et al.) and ceramics (Rueschhoff et al., 2016; Costakis Jr et al., 2016). In these applications, 3D geometry is printed in the liquid state, relying on fluid yield stress to form a self-supporting structure. After printing, the entire structure is cured or fired, during which the part experiences significant shrinkage. This secondary curing process can be lengthy (16 hours in the case of Pierson et al.), so it is important to have confidence in the as-printed geometry, which is scaled by the shrinkage factor. Of course, implicit is the assumption of homogeneous shrinkage, which is valid in the example applications.

A fair assessment of LOS inspection also requires an examination of its disadvantages. The output of the LOS methods are distances between LOS distributions, which are in turn based on D2 distance, D3 area, D4 volume, or A3 angle shape vectors. This lacks the clear physical interpretation of pointwise physical distance to a true surface provided by conventional methods, although LOS inspection can offer more high-level insight. The advantages associated with the high-level nature of the method are tempered by the fact that LOS inspection also does not provide
the same level of detail as pointwise distance measurements. Small local anomalies in otherwise accurate features could be lost in the noise, and it does not provide the user with information as to what is wrong with features that are flagged as inaccurate. This can be mitigated somewhat via finer feature segmentation, but at higher computational cost. Finally, LOS inspection relies on 3D scanning technology that has yet to match the precision and reliability of CMM. This gap is expected to gradually close, though, as the technology advances.

These advantages and disadvantages suggest an application-specific tradeoff between LOS inspection and more conventional methods, but it is not necessarily an either-or choice. LOS inspection could be used in combination with conventional methods to provide supplemental information. For 3D scanned parts of similar scale, the same point clouds could be analyzed using both LOS inspection and pointwise deviation to yield a more complete evaluation of as-built geometry. LOS inspection could also be used as a screening tool in conjunction with CMM/GD&T. In this role, LOS inspection could screen for suspected bad parts, thereby reducing the number of parts sent to CMM. It could also identify specific features of concern, the focusing CMM inspection resources on the features most likely to be out of tolerance.

2.7 Conclusion

This work proposes a point-cloud-based inspection method based on the novel LOS distribution, which is a generalization of the shape distribution concept used in object recognition. LOS distributions are probability density functions that encode both shape and pose information based on geometric relationships between points in two point clouds. The LOS distribution between features of interest is constructed from both the scanned artifact and the nominal model, and the difference between the distributions is used as a quality metric. Inspection via LOS distributions permits feature-based inspection with respect to form, position, and orientation with or without datum surfaces. Although the method can be applied to parametric geometries in traditional manufacturing, it is particularly well suited to inspection of nonparametric geometries in the high-mix, low-volume production environment typical of AM. The method is tested on several physical AM
builds, and the results show that the method is able to reliably differentiate between acceptable and unacceptable features.

References


3 Multi-resolution In-plane Alignment and Error Quantification

Additive manufacturing has shown its capability in producing complex geometries. Due to the additive nature, the in-situ layer-wise inspection of geometric accuracy is essential to making AM reach its full potential. This work proposes a novel automated in-plane alignment and error quantification framework to distinguish the fabrication, measurement, and alignment errors in additive manufacturing. In this work, a multi-resolution framework based on wavelet decomposition is proposed to automatically align two-dimensional point clouds via a polar coordinate representation and then to differentiate errors from different sources based on a randomized complete block design (RCBD) approach. In addition, a two-step optimization model is proposed to find the best configuration of the multi-resolution framework. The proposed framework can not only distinguish errors attributed to different sources but also evaluate the performance and consistency of alignment results under different levels of details. A sample part with different featured layers, including a simple free-form layer, a defective layer, and a layer with internal features, is used to illustrate the effectiveness and efficiency of the proposed framework. The proposed alignment method outperforms the widely used Iterative Closest Point algorithm. This work fills a research gap of state-of-the-art studies by automatically quantifying different types of error inherent in manufacturing, measuring, and part alignment.

3.1 Introduction

Additive manufacturing (AM) has shown its significant advantages over traditional subtractive and forming manufacturing processes for offering agile production of complex geometries (Gibson et al., 2014). Nevertheless, to make AM reach its full potential, geometric quality inspection is indispensable, especially in high-value applications such as biological structure substitutes in medical areas (Jin et al., 2015) and airfoils in aerospace industry (Liu et al., 2017) where component failures are not tolerated. Due to the layer-wise additive nature of AM processes, the layer-by-layer in-situ inspection and monitoring (rather than offline quality inspection) are desired in resolving failures on the current layer or out-of-control process variation as soon as they are recognized before the
errors get accumulated and affect the subsequent layers.

The geometric accuracy of each layer is usually referred to as in-plane (x-y plane) shape deviation (Cheng et al., 2018; Wang et al., 2017; Luan and Huang, 2017). Usually, the shape of the current layer can be described by tracing thousands of points along the specified boundary using optical/vision measurement equipment or by slicing a thin layer of 3D point cloud obtained from 3D scanning. The output of measurement can be considered as a two-dimensional point cloud, which must be aligned with the corresponding base-truth CAD model prior to evaluating the geometric accuracy of the layer. Common alignment methods coming with a point cloud processing software package rely on least-square fitting algorithms to find the optimal rigid transformation that minimizes the sum of squared distances between the measured alignment points and their target reference points. However, the performance of least-square fitting in aligning two-dimensional point clouds is worth to be examined. Moreover, in practice, the dimensional accuracy is associated with datum features from a Geometric Dimensioning and Tolerancing (GD&T) perspective, but the least-square fitting methods are not able to automatically align point clouds based on such datum features without first performing a global rough alignment. A phase-correction–based alignment method is proposed by Jin et al. (2019) for datum-based alignment. This method is adopted and improved in this work to align a point cloud with different levels of details maintained, and to evaluate the alignment performance in terms of alignment error and consistency.

Based on the aligned point cloud, the point-wise deviation, as a common metric for evaluating dimensional accuracy, is calculated for the currently inspected layer. It is usually presented with a deviation color map (e.g., Rao et al., 2016) or a deviation profile or using polar coordinates (e.g., Song et al., 2014). It is worth pointing out that the variation observed in the point-wise deviation is the result of systematic error, measurement error, and alignment error. The systematic error mainly results from a persistent issue, such as an inappropriate design or out-of-control fabrication process parameter, and leads to a consistent error, which is the one to trigger an alarm informing engineers to remove the error during in-situ monitoring. The measurement error is usually referred to as random error, which is unavoidable due to the limited resolution or precision of the mea-
suring equipment. The alignment error provides a measure of how well a particular point cloud alignment/registration method performs in aligning the sample layer to the nominal model. The performance of alignment might not only be affected by the algorithm itself, but also varies by using different levels of detailed point cloud information. The systematic error and random error can usually be differentiated by measuring the sample multiple times and then using Analysis of Variance (ANOVA) or Average and Range method. However, there is a gap in existing research regarding how best to identify the alignment error and distinguish three error terms by an instant scan during in-situ monitoring. In this work, a detailed review is provided to summarize the existing deviation modeling and error quantification methods in additive manufacturing along with common point cloud alignment methodologies.

In addition, after resolving the impact of alignment and random errors, the change point of systematic error distribution is desired to be identified as early as possible so that engineers can be informed to resolve the errors from the change point. Change point is the time instant when the distribution of a random process changes. In the in-situ inspection and monitoring of AM processes, the systematic error calculated based on the multi-resolution framework provides real-time information about when (at which layer) a systematic deviation occurs. This problem can be solved via quickest change detection (QCD), which is often applied in detecting changes in a time series as quickly as possible under a probability of false alarm (PFA) constraint. Usually, the distribution before the change point can be obtained based on the normal operating condition of the system. However, the post-change distribution might not be readily available due to the unexpected nature of the fabrication variations. Existing QCD methods can be roughly classified into Bayesian and Minimax (non-Bayesian) methods. The Bayesian procedures are applied when the prior knowledge of the change point is available. A well-known procedure is the Shiryaev method (Shiryaev, 1963) which aims at minimizing the average detection delay (ADD) subject to an upper bound on the probability of failure (POF). If the prior probability of change point is unknown, the non-Bayesian methods, such as the cumulative sum (CUSUM) (Page, 1954) method and the Shiryaev-Roberts (SR) procedure (Roberts, 1966), are applied to minimize the delay considering
a lower bound of Average Run Length (ARL). The limitation of these works is that only a bi-
nary hypothesis test (single pre-change and post-change model) is considered. In Lai (1998) and Tartakovsky (2004), methods of the non-Bayesian QCD formulation with multiple post-change models have been discussed. However, those studies on QCD did not consider the objective of ADD for QCD problems with multiple post-change models under a non-Bayesian setting. In this work, a QCD method for minimizing the ADD (Nath and Wu, 2018) under a non-Bayesian setting is integrated with the multi-resolution framework to accomplish a more comprehensive online inspection scheme for AM processes.

Regarding methodological contributions, this work proposes a multi-resolution framework for automatically aligning a layer to its nominal shape with different levels of details and distinguishing the errors due to fabrication, measurement, and alignment. In addition, this research is among the first attempts that provide an efficient and effective solution to find the change point (layer) in layer-by-layer inspection of an AM process with regard to in-plane systematic deviation. Shapes are modeled using a polar coordinate representation/profile, which is decomposed into different levels of components via wavelet decomposition. Then, an alignment method based on phase correction is conducted on different wavelet components. After alignment, the randomized complete block design (RCBD) approach (Montgomery, 2013) is applied to quantify the manufacturing (systematic) error, alignment error, and random error based on the deviation profile. In addition, a two-step optimization model is proposed to optimize the efficiency and effectiveness of the proposed framework. This work mainly focuses on illustrating the multi-resolution framework; therefore, the procedure of datum-based alignment, which has been elaborated upon in Jin et al. (2019), will not be repeated in this work. Three layers of a 3D printed human heart are considered as examples to demonstrate the effectiveness of proposed methods. The first two layers are designed to show the procedure of newly proposed multi-resolution framework. The third layer is chosen as an example for showing capability of the proposed method with datum-based alignment to deal with complex layers which have internal or concave features.
3.2 Literature Review

3.2.1 Point Cloud Alignment

The goal of an alignment or registration process is to transform a model from the internal coordinate system of a measurement device into a global coordinate system that the user desires. Usually, it keeps the reference (nominal) model fixed and lets the source point cloud obtained from a 3D scanner to best match the reference by certain transformation. Existing registration algorithms can be classified coarsely into rigid and non-rigid approaches. Rigid approaches assume a rigid environment such that the transformation can be modeled using only 6 degrees of freedom (DOF). Non-rigid methods, on the other hand, allow for a higher number of DOF in order to cope with non-linear or partial stretching or shrinking of an object. Because alignment is a pre-processing step in quality inspection, rigid approaches are often preferred for aligning the scanned point cloud to the reference model. Detailed reviews of rigid registration are provided by Bellekens et al. (2014), Bellekens et al. (2015), Tam et al. (2013), and Markelj et al. (2012). In the literature, most applications employ either a simple Singular Value Decomposition (SVD) or Principal Component Analysis (PCA) for registration, or use a more advanced iterative scheme based on the Iterative Closest Point (ICP) algorithm. In general, most rigid alignment methods, including ICP, rely on least-squares fitting algorithms to find the optimal rigid transformation. In ICP, the error of alignment (or the distance between the two point clouds) is minimized by iteratively revising the transformation matrix that aligns the two point clouds.

Unfortunately, limited work has been done on how these registration methods work under GD&T specifications. Some commercial software/hardware providers (e.g., BuildIT Construction, 2019) offer alignment options based on a given feature/datum, which is manually specified by users. Since specific alignment algorithms used by a particular software/hardware provider are usually confidential, useful information is difficult to obtain online. Indeed, the most popular alignment options include the best fit alignment, feature-based alignment, and degrees-of-freedom alignment (Woodward et al., 2018). Among the three, the feature-based alignment can
align two point clouds based on some important point sets specified by users, but it requires a global least-squares fitting before performing feature-based local adjustment. In practice, it is possible that defects on the sample part might deteriorate the alignment accuracy and cause point cloud misalignment, which might impact the accuracy of local adjustment and establish an even worse alignment.

The alignment methods proposed in this work are motivated by the phase correction approach (Wolberg et al., 2000) for image registration. The main idea of phase correction is to find the phase shift by representing the original image in the polar coordinate system. The proposed methods have been initially studied by Jin et al. (2019), which can directly align the reference shape to the sample point cloud without conducting a rough pre-alignment or manual intervention. In this work, the alignment idea is advanced by using a multi-resolution alignment methodology based on the wavelet components to verify the alignment consistency under different levels of resolution. The wavelet decomposition is a common technique that applied in signal and image processing to investigate the information with different level of details. For point cloud data, this technique is usually applied for point cloud denosing (Huang et al., 2009, Rosman et al., 2013), and surface decomposition/reconstruction (Manson, et al., 2008, Chen et al., 2011, Jakovljevic et al., 2015). Torre-Ferrero et al. (2008) proposed a 3D registration method based on wavelet decomposition. However, the wavelet decomposition is applied on the circon image, which is a shape presentation for 3D point cloud, rather than on the point cloud directly. Also, it did not consider the alignment at different resolutions. To the best of our knowledge, this is the first attempt in solving such complex alignment problems under a multi-resolution framework to hierarchically evaluate the alignment performance and analyze the variation from different sources.

### 3.2.2 Deviation Modeling and Quantification in AM processes

In previous studies, most shape deviation modeling approaches are proposed under the context of design compensation by either changing the CAD, STL, or slicing techniques.

Tong et al. (2003) and Tong et al. (2008) proposed parametric error models to map the effects
of all errors in STL and FDM processes into “virtual” parameters without explicit meanings. The coefficients of the models that are considered as a global measure of accuracy can be estimated through regression of the measurement data represented in a Cartesian Coordinate System (CCS).

In a different direction, a series of works (Huang et al., 2014, Song et al., 2014, Huang et al., 2015, Sabbaghi et al., 2016, Wang et al., 2017) on predictive modeling and compensation have been conducted by modeling the deviation in a Polar Coordinate System (PCS). In this representation, the Cartesian coordinates \((x, y)\) of a manufactured shape are transformed into polar coordinates \((\theta, r(\theta))\), where \(\theta\) is the angle of a point as defined with respect to coordinate axes printed on the product, and \(r(\theta)\) is the observed radius for \(\theta\). The desired shape is represented by the nominal radius set \(r_0(\cdot)\). The deviation at \(\theta \in [0, 2\pi]\) is then defined as

\[
\Delta r(\theta, r_0(\cdot)) = r(\theta) - r_0(\theta)
\]

This deviation modeling approach is adopted in this work. In particular, the shape profile represented with the polar coordinates are considered as an observed “signal” of each layer for wavelet analysis.

Different from modeling error based on experimental data, other kinds of error modeling approaches in AM (e.g., Navangul et al., 2013; Zha et al., 2015) are proposed to model the chordal error from the tessellation, which is computed as the distance between the STL facet and the CAD surface. In medical applications (Arrieta et al., 2012, Pinto et al., 2015), the geometric error is evaluated both globally using volumetric intersection indexes calculated over segmented Computed Tomography scans and locally by applying a 3D surface map with a color code that allows for differentiating regions for which the model is overestimated, underestimated, or correctly estimated.

In contrast, relatively few studies have been conducted on how to distinguish the errors from different sources. In Song et al. (2014), the errors are categorized into the extruder positioning error before material deposition and the processing error occurred during the printing process or
after material deposition. Zhu et al. (2017) proposed a deviation modeling approach from a Design for AM (DfAM) perspective. In their work, the systematic deviation is considered as the in-plane deviation between the actual printed shape and the nominal shape of a 2D layer, and the random errors are generated by translating the contour points with a given distance derived from a random field theory. However, none of these studies have considered how to differentiate the systematic, alignment, and random errors from the deviation profile generated by a single scan of a layer. The methodological contribution of this work is to propose a novel multi-resolution alignment and inspection framework to fill this research gap.

3.3 Method

The proposed automatic multi-resolution alignment and inspection framework is illustrated in Figure 3.1. The sample point cloud is obtained by 3D scanning while the reference (nominal) point cloud is generated from the STL file. First, both the sample point cloud and reference (nominal) point cloud of a given layer are represented in a shape profile with polar coordinates. Then, \( \mathcal{L} \)-
level wavelet decomposition is conducted on the two profiles, respectively, to obtain $L$ levels of
detail for each profile. Next, each level of detail coefficients of the sample profile is aligned to the
corresponding detail coefficients of the reference profile by finding the optimal phase shift among
the $N$ best candidates. Note that the proposed alignment methods are proposed based on the
assumption that the sample point cloud and reference point cloud have already been aligned along
the z-axis. The randomized complete block design (RCBD) is then conducted on the point-wise
deviation calculated based on different levels of alignment results to distinguish the errors from
different sources. To improve the alignment effectiveness and efficiency, a two-step optimization
model is proposed to find the optimal level of wavelet decomposition and the maximum number
of candidates to be used in the alignment process. Based on the systematic error obtained from
RCBD, a QCD method is implemented to identify the layer where the earliest systematic error
distribution changes during the printing process. The details of these steps will be explained next.

3.3.1 Shape Profile

A shape profile is a piecewise polynomial function, $r(\theta)$, estimated based on the polar coordinates
(Jin et al., 2019). Given a point $(x, y)$ from the point cloud, the corresponding polar coordinates
$(\theta, r)$ can be expressed as

$$\theta = \text{atan2}(y - y_0, x - x_0)$$

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

(3.2)

where $(x_0, y_0)$ is the centroid of the 2D shape. The angle $\theta$ of a point should be within $[-\pi, \pi]$, and its corresponding radius should be normalized into a range of $[0, 1]$ to eliminate scaling
effects. Based on the angles and radius obtained from the point clouds, the corresponding piecewise
polynomial function can be estimated. An example is provided in Figure 3.2 to show the repre-
sentations of a 2D point cloud in the CCS, PCS, and shape profile. In order to establish future
point-wise correspondence between the reference point cloud and sample point cloud, the same
query set $\Theta$ is applied on both reference and sample profiles to generate the corresponding radii.
For any $\theta \in \Theta$, the reference radius is denoted as $r_0(\theta)$ and the observed radius in the sample is
denoted as \( r(\theta) \). The discretized shape profiles using the query set \( \Theta \) are denoted as \( r^0_j \) and \( r_j \) for the reference model and sample model, respectively. The reason for applying the same query set is to ensure same number of points are sampled with the same resolution in both sample shape profile and reference shape profile.

![Figure 3.2: An example of shape representation in CCS, PCS, and shape profile.](image)

### 3.3.2 Wavelet Decomposition

Wavelet decomposition (Mallat, 1989) is a common technique usually applied in signal or image processing to learn the details of information at different resolutions. Like a signal or image, point cloud data also contains local variations at different spatial frequencies. With a multi-resolution representation using wavelet decomposition, the shape profile can be interpreted hierarchically with different levels of detail. In this work, a shape profile can be considered as a “signal”, in which angle is analogous to time, and radius is analogous to amplitude. Therefore, the original shape profile \( r_j \), as shown in Figure 3.3, can be decomposed into an approximation coefficient vector \( r_{j-1} \) and a detail coefficient vector \( d_{j-1} \) by down-sampling the signal after passing through a high-pass filter \( H_{ID} \) and a low-pass filter \( L_{OD} \), respectively. The approximation coefficients are a lower resolution representation of the original profile, and the detail coefficients are defined as the difference of information between the original profile and approximation at resolution \( j - 1 \).
Technically, the two wavelet components can be obtained by

\[
\begin{align*}
  r_j(\theta) &= \sum_{k \in \mathbb{Z}} r_{j-1}(k) \phi_{j,k}(\theta - k) + \sum_{k \in \mathbb{Z}} d_{j-1}(k) \psi_{j,k}(\theta) \\
  r_{j-1}(k) &= \langle r_j(\theta) \phi_{j,k}(\theta) \rangle, \quad d_{j-1}(k) = \langle r_j(\theta), \psi_{j,k}(\theta) \rangle
\end{align*}
\]  

(3.3)

where \(\phi_{j,k}(\theta - k)\) is the scaling function, and \(\psi_{j,k}(\theta)\) is the wavelet function. The approximation \(r_{j-1}\) can be further decomposed into two lower level components by following the similar procedure. It is worth noting that the index \(j\) or \(j - 1\) denotes only the detail level the coefficient vector contains. For instance, if 4-level wavelet decomposition is performed on a shape profile \(r_j\), then the level 1 to level 4 detail coefficient vectors are denoted as \(d_{j-1}, d_{j-2}, d_{j-3},\) and \(d_{j-4}\).

Figure 3.3: Wavelet decomposition of shape profile.

Considering the tradeoff between the alignment performance and computational efficiency, both wavelet levels and wavelet functions need to be optimized. In practice, the maximum level of wavelet decomposition can be determined using

\[
\mathcal{L}_{\text{max}} = \left\lfloor \log_2 \left( \frac{n}{n_w} - 1 \right) \right\rfloor
\]  

(3.4)

where \(n\) is the length of the signal, \(n_w\) gives the length of the decomposition filter associated with the selected mother wavelet. It is worth pointing out that the optimal \(\mathcal{L} \in \{1, 2, \ldots, \mathcal{L}_{\text{max}}\}\) can be obtained using an optimization model illustrated in Section 3.3.5.
3.3.3 Multi-Resolution Alignment based on Coefficient Profile

The detail coefficient vector obtained from the wavelet decomposition is named as the coefficient profile. If a \( L \)-level wavelet decomposition is performed on the original profile, there are \( L \) coefficient profiles to represent the original profile with different spatial frequencies. Let \( d^0_{j-l} \) denotes the nominal coefficient profile while \( d_{j-l} \) represents the sample coefficient profile at \( l^{th} \) level of resolution. The main idea of alignment based on the coefficient profile is to find the optimal phase shift for the sample coefficient profile \( d_{j-l} \) in the corresponding reference coefficient profile \( d^0_{j-l} \) at each level \( l \), \( l \in \{1,2,\ldots,L\} \). The alignment based on the sample and reference coefficient profiles at different resolutions can be accomplished in a three-step procedure.

First, for each level of sample coefficient profile \( d_{j-l} \), find \( N \) best matches in its corresponding extended reference coefficients profile \( d^0_{j-l} \) by similarity search (Hetland, 2009), and calculate the corresponding \( N \) phase shifts. The phase shift at level \( l \) based on the \( i^{th} \) best match is denoted as \( \Delta \theta_{li}, l \in \{1,2,\ldots,L\} \) and \( i \in \{1,2,\ldots,N\} \). It is worth noting that the reference coefficient profile is extended with one cycle on the left and one cycle on the right to ensure that at least one best match can be found for the sample coefficients profile. An example of finding 5 best matches in the 2nd level of the sample coefficient profile of the example shape from Figure 3.2 is provided in Figure 3.4. The horizontal axis shown in Figure 3.4 represent the indices of \( \theta \)s. The corresponding \( \theta \) values can be obtained from the query set with a given index. Each of the 5 matches has a phase shift, which indicates how much the sample coefficient profile needs to be moved along the \( \theta \) axis. A positive \( \theta \) indicates a movement to the right, while a negative value indicates a shift to the left.

Second, for each level \( l \), find the optimal phase shift \( \Delta \theta^*_l \) among all the candidates. Specifically, each \( \Delta \theta_{li} \in \Delta \theta(l) \) is evaluated by calculating the total absolute point-wise deviations based on the aligned sample and reference shape profiles using \( \Delta \theta_{li} \). The optimal phase shift \( \Delta \theta^*_l \) at the \( l^{th} \) level of resolution is determined by

\[
\Delta \theta^*_l = \arg\min_{j \in \{1,2,\ldots,N\}} \sum_{k=1}^{n} |\Delta \rho^{(l,j)}(\theta_k)| \quad \text{for} \quad \forall i \in \{1,2,\ldots,L\}
\] (3.5)
where $\Delta r^{(l,j)}(\theta_k)$ is the deviation at angle $\theta_k \in \Theta$ based on the phase shift $\Delta \theta_{l1}$, and $n$ is the length of the query set $\Theta$. Continuing with the example shown in the first step, the evaluation of each candidate shift in $\Delta \theta_2$ is shown in Figure 3.5 along with the aligned shape profiles. Among all shift values in $\Delta \theta_2$, $\Delta \theta_{23}$ and $\Delta \theta_{24}$ achieve the smallest error among the 5 candidates. In this case, we can arbitrarily select either $\Delta \theta_{23}$ or $\Delta \theta_{24}$ as the optimal phase shift for the 2nd level of alignment.

Third, align the $l^{th}$ level of coefficient profiles $(d_{j-l}^0, d_{j-l})$ and original shape profiles $(r_j^0, r_j)$ using the optimal phase shift $\Delta \theta_{l}^{*}$. Based on the optimal phase shift of each level obtained from the previous step, the 4-level multi-resolution alignment result of the example shape is shown in Figure 3.6.

Similar alignment procedures can be conducted based only on the segments associated with datum features or a collection of user-defined features. The detail steps and applications are elaborated upon in Jin et al. (2019).

### 3.3.4 Deviation quantification

According to the alignment shape profiles at each level $l$, the deviation at each point (angle) $k$ can be calculated and denoted as $\nu_{lk}$. One challenge of deviation quantification using the proposed
Figure 3.5: Find the optimal phase shift for the 2nd level of the coefficient profile.

\[
\Delta \theta_{21} = -5.215 \quad \text{Dev}=201.814
\]

\[
\Delta \theta_{22} = 1.068 \quad \text{Dev}=201.814
\]

\[
\Delta \theta_{23} = -4.668 \quad \text{Dev}=38.379
\]

\[
\Delta \theta_{24} = 1.615 \quad \text{Dev}=38.379
\]

\[
\Delta \theta_{25} = -4.939 \quad \text{Dev}=112.657
\]

Arbitrarily select one as \( \Delta \theta'_2 \)

Figure 3.6: Multi-resolution alignment of detail coefficient profiles and shape profiles.

multi-resolution alignment approach is in checking the consistency of quantification results at different levels. In particular, it is likely that the values of deviation at different locations (i.e., angles) are quite different. To resolve this issue by “noise reduction” (Montgomery, 2013), RCBD is ap-
plied on the point-wise deviation matrix to learn the sources of variance. In this work, each level of alignment is considered as a treatment while each sample point (angle) is considered as a block as demonstrated in Table 3.1. Given the point cloud with size n, m points are sampled from the n points. The m sample points can be points associated with the critical features, or can be obtained via uniform sampling to eliminate bias.

Table 3.1: Point-wise Deviation Based on the Multi-resolution Alignment

<table>
<thead>
<tr>
<th>Sample point</th>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
<th>...</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment 1</td>
<td>1</td>
<td>v_{11}</td>
<td>v_{12}</td>
<td>v_{13}</td>
<td>...</td>
</tr>
<tr>
<td>Treatment 2</td>
<td>2</td>
<td>v_{21}</td>
<td>v_{22}</td>
<td>v_{23}</td>
<td>...</td>
</tr>
<tr>
<td>Treatment 3</td>
<td>3</td>
<td>v_{31}</td>
<td>v_{32}</td>
<td>v_{33}</td>
<td>...</td>
</tr>
<tr>
<td>Treatment L</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>\mu_v^1</td>
<td>\mu_v^2</td>
<td>\mu_v^3</td>
<td>\mu_v^m</td>
<td>\mu_v^L</td>
</tr>
</tbody>
</table>

To check the consistency of alignment at different resolutions, a hypothesis test shown in Eq. 3.6 is conducted on the mean values of L treatments (rows), \( \mu_v^1, \mu_v^2, \ldots, \mu_v^L \).

\[
H_0 : \mu_v^1 = \mu_v^2 = \ldots = \mu_v^L.
\]

\[
H_1 : \text{at least one “equal” is not satisfied}
\]

If the alignment results of different levels are consistent, a p-value greater than the pre-determined significant level should be observed (i.e., fail to reject the null hypothesis). Inconsistency of alignment would indicate that the deviation measured in Table 3.1 not only results from the systematic or measurement error but from the alignment. Therefore, more candidates (best matches) should be evaluated in the alignment procedure. The stop criterion of stopping adding candidates is elaborated in section 3.3.5. If the inconsistency still exist after the stop criterion is satisfied, then it is...
high likely that the defects on the layer has affected the alignment accuracy. As a result, a large systematic error and alignment error might be observed in the result.

Moreover, by following the procedure of ANOVA, the general systematic error, the alignment error at the \(i^{th}\) level, and random error at each sample point can be calculated:

**Systematic error:**
\[
e_{s} = \mu_v = \frac{\sum_{i=1}^{L} \sum_{j=1}^{m} v_{ij}}{L m}. \tag{3.7}
\]

**Alignment error for level \(i\):**
\[
e_{a}^{(i)} = \mu_i - \mu_v, \quad i \in \{1, 2, ..., L\}. \tag{3.8}
\]

**Random error for each point of level \(i\) at block \(j\):**
\[
e_{r}^{(i,j)} = v_{ij} - \mu_i, \quad i \in \{1, 2, ..., L\}, \quad j \in \{1, 2, ..., m\}. \tag{3.9}
\]

### 3.3.5 Optimal setting for point cloud alignment

As illustrated in section 3.3.3, the performance of alignment can be improved by searching the optimal level of wavelet decomposition \(L\) and the maximum number of best phase shifts \(N\) when aligning the sample and reference coefficient profiles. Given \(N\) and \(L\) are values, the optimization model of the first step for determining the best phase shift at each level can be expressed as

\[
z_1(N, L) = \min \sum_{l \leq L} \sum_{i \leq N} \left( \sum_{\theta \in \Theta_s} \Delta r^{(l,i)}(\theta_k) \right) x_{li} \tag{3.10}
\]

**Subject to:**
\[
\sum_{i} x_{li} = 1, \quad \forall l \in \{1, 2, ..., L\} \tag{3.11}
\]
\[
x_{li} \in \{0, 1\}, \quad \forall l \in \{1, 2, ..., L\}, \forall i \in \{1, 2, ..., N\} \tag{3.12}
\]

where \(x_{li}\) is the decision variable (\(x_{li} = 1\), if the \(l^{th}\) level of decomposition uses \(i^{th}\) match; \(x_{li} = 0\), otherwise). The total point-wise deviation \(\sum_{\theta \in \Theta_s} \Delta r^{(i,j)}(\theta)\) depends on the sample points set \(\Theta_s\). With different sampling strategies, the objective value may vary. The second-step optimization is to minimize the sum of the objective value from the first step and the computational cost to
determine the optimal values of \( \mathcal{L} \) and \( \mathcal{N} \).

\[
\min \alpha z_1(\mathcal{N}, \mathcal{L}) + \beta(\mathcal{N}\tau + \mathcal{L}\omega) \quad (3.13)
\]

\[
s.t. \quad \mathcal{N}, \mathcal{L} \in \mathbb{Z}^+, \quad 3 \leq \mathcal{L} \leq \mathcal{L}_{\text{max}} \quad (3.14)
\]

Note that the decision variables \( \mathcal{N} \) and \( \mathcal{L} \) are integers, where \( \mathcal{L} \) should at least be 3 to accomplish the multi-resolution framework and be less than \( \mathcal{L}_{\text{max}} \) given by Eq 3.4. The unit computation costs for wavelet decomposition and alignment are denoted as \( \omega \) and \( \tau \), respectively. The penalty coefficients \( \alpha \) and \( \beta \) should be adjusted by considering the magnitude of deviation. In this work, the above two-step optimization model is solved by a heuristic algorithm. For each pair of values of \( \mathcal{N} \) and \( \mathcal{L} \), the integer programming problem \( z_1(\mathcal{N}, \mathcal{L}) \) is solved directly. If the improvement of the objective value is less than the defined threshold \( \delta \), the algorithm stops and reports the solution; otherwise, another iteration is performed.

3.3.6 Quickest Change Point Detection

Given a specific decision rule \( \delta \) for determining the change point, the QCD problem based on the systematic error can be formulated as:

\[
\min ADD(\delta)
\]
\[ s.t. \ PFA(\delta) \leq \alpha \] (3.15)

where \( PFA(\delta) = P(\hat{\delta} < l | R_n^X) \) and \( ADD(\delta) = E[\hat{\delta} - l | \hat{\delta} > l] \). Here we denote the true change point (layer) as \( l \) and the estimated change point as \( \hat{l} \) based on the decision rule \( \delta \). The mapping from an observation to the change is denoted as \( R_n^X \), for which \( \delta(R_n^X) = \hat{l} \).

Let \( X_n, n = 1, 2, \ldots \) be a sequence of observed systematic errors at moment \( n \) until the printing stops. Then, the probability density function (pdf) before \( (n < l) \) is denoted as \( f_{0,n}(X_n|X^{1:n-1}) \), and let the pdf after the change point \( (n > l) \) be \( f_{i,n}(X_n|X^{1:n-1}) \) based on the \( i^{th} \) post-change model where \( i = 1, 2, \ldots, M \) with \( M < \infty \). A typical change point scenario is demonstrated in Figure 4. At any moment \( n \), a decision is made based on the following hypothesis test:

\[
H_0 : l < n \\
H_1 : l \geq n
\] (3.16)

\[ H_0 : l < n \\
H_1 : l \geq n \]

Figure 3.7: Demonstration of change point in a time series.

In practice, due to unpredictable variations in fabrication, both the prior probability of change point \( P(l = k) = \pi_k, k = 1, 2, \ldots \) and the post-change models are unknown. Based on the derivation in Page (1954), Lai (1998), and Nath and Wu (2018), the change point can be identified according to the following theorem.

**Theorem 1.** For a given \( PFA \) upper bound \( \alpha \), the change point is detected using \( \hat{l} = \inf \{ n > l : \Lambda(n) \geq \frac{M\hat{\mu}}{\alpha} \} \), where \( \Lambda(n) = \sum_{i=1}^{M} \Lambda_i(n) \) is the sum of the Shiryaev-Roberts (SR) statistic.
Λ_i(n), and \( \bar{l} = \sum_{k=1}^{\infty} k \pi_k \) is the mean over the prior distribution of change point.

The SR statistic \( \Lambda_i(n) \) for detecting a change corresponding to the \( i^{th} \) post-change model can be calculated by:

\[
\Lambda_i(n) = \sum_{k=1}^{n} \exp[Z_i^{k:n}]
\]

where

\[
Z_i^{k:n} = \sum_{t=k}^{n} \log \frac{f_i,t(X_t|X_1:t-1)}{f_0,t(X_t|X_1:t-1)}
\]

which can be obtained by the likelihood ratio test on the given hypotheses. The detection rule given in Theorem 1 is an extension of SR procedure, which is adapted to detect a change point with multiple post-change models (Nath and Wu, 2018). Accordingly, the upper bound of false alarm rate is \( \alpha \bar{l} \). Based on the theorem, the mean \( \bar{l} \) can be calculated based on the prior distribution of all possible positions \( k, k = 1, 2, \ldots \) In the non-Bayesian setting, the prior probability of the change point \( \pi_k \) is unknown. However, it can be estimated by:

\[
\pi_k = \frac{h_0(X_k)}{\sum_{k=1}^{n} h_0(X_k)}
\]

where \( h_0(X_k) \) is the probability that \( X_k \) belongs to the empirical distribution of systematic error measured before moment \( n \) under the normal system condition.

A simulated example using a time series of 100 observations following a normal distribution is provided in Figure 3.8 to illustrate the QCD process. It is assumed that the distribution change happens at the 51st observation. The observation before the change point follows the normal distribution \( N(0, 0.1^2) \) while the observations after the change point are from the normal distribution \( N(0.4, 0.2^2) \). The QCD is performed by considering a single post-change model and multiple post-change models, respectively. The tolerance \( \alpha \) on the PFA is set to be 0.05.

In the case with the single post-change model, it is assumed that the post-change model is following \( N(0.4, 0.2^2) \). With this single post-change model, \( M = 1 \) and the detection statistic can be calculated directly from Eq.(3.17). The result of QCD based on the single post-change is demonstrated in Figure 3.9. It is observed that the detection statistic exceeds the threshold,
Figure 3.8: Overview of simulation data.

\[ \bar{Ml}_\alpha = \bar{l} = 1.004 \times 10^5, \] at the 53rd observation. Therefore, in this case, the detection delay is equal to 2.

Figure 3.9: The QCD result based on single post-change model.

When considering multiple post-change models (M=3) following \( N(0.38, 0.22^2), N(0.42, 0.18^2) \) and \( N(0.4, 0.2^2) \), a similar detection procedure can be performed based on Theorem 1. Note that
two candidate post-change models are slightly different from the true post-change distribution to simulate the inaccuracy of parameter estimation in practice. The result of considering multiple post-change models is presented in Figure 3.10 that shows the same estimation of change point.

Figure 3.10: The QCD result based on multiple post-change model.

To further investigate the sensitivity of the QCD method, numerical experiments are conducted by considering different distribution parameters for the single post-change model. Specifically, given $N(0, \sigma^2)$ as the distribution before the change point, the data after the change point is generated from $N(1, \sigma^2)$ with $\sigma \in [0.01, 2]$. The ADD is measured and plotted against $\sigma$ in Figure 3.11 by considering the single post-change model following $N(1, \sigma^2)$ as well. It is observed that the performance of QCD is much more unstable with larger ADD variation as the variance of the data increases. It can be explained by the fact that with a large variance, the distribution before and after the change point cannot be well separated, and therefore, the likelihood ratio test might not work effectively in calculating the SR statistics. The variation of ADD can also be observed when changing the distribution parameters of the post-change model as shown in Figure 3.12.
3.4 Experiment

To validate the proposed framework, a human heart shown in Figure 3.13(a) was printed with MakerGear FFF machine with 1.75mm PLA filament (see Figure 3.13(b)), and scanned by Nikon X TH 225 ST µCT scanner (see Figure 3.13(c)). The output of CT scanning was a TIFF image stack, which was converted to a point cloud by ImageJ (see Figure 3.13(d)). The proposed multi-
resolution framework is implemented with Matlab on a desktop with Intel core i7 CPU.

In this work, the layer-wise in-plane deviation inspection is simulated by observing X-Y coordinates of a thin slice of the 3D point cloud. It is assumed that the printed sample and the nominal model have been aligned along the z-axis. The layer thickness values of the nominal point cloud and the scanned point cloud are determined based on their scales to ensure that both models have the same number of layers.

![Figure 3.13: Experiment setup.](image)

In the experiment, three sample layers (see Figure 3.14) are to be inspected using the multi-resolution framework. The 1st layer is a simple conforming free-form layer used for comparing the proposed alignment method to the default ICP methods provided in Matlab. The 2nd layer is a defective layer used for demonstrating the effectiveness of the error quantification method. The 3rd layer has complex internal features showing the positions of key tissues of a human heart. This layer is chosen to illustrate how to accommodate our methods on inspecting layers without convexity constraints. Moreover, an additional experiment is conducted on 100 layers from the upper half of the human heart for illustrating the QCD of systematic error.
3.4.1 The 1st layer — a simple layer

The 1st layer is known as a conforming layer, whose shape profile along with its corresponding reference shape profile are shown in Figure 3.15. To achieve better alignment, the alignment parameters, \( N \) and \( L \), are first optimized using Algorithm 1. The upper bound of wavelet decomposition level \( L_{\text{max}} \) is 8 for this shape profile. The penalty coefficient \( \alpha \) of the objective value and \( \beta \) of the computational cost are equal to 2 and 0.1, respectively, in this experiment. The result of the optimization is shown in Figure 3.16, suggesting that 4-level wavelet decomposition should be conducted on the original shape profiles. For each level of alignment, the best phase shift for aligning the sample coefficient profile to the reference coefficient profile is found among 25 candidates. Then, the coefficient profiles, as well as the shape profiles, can be aligned based on the optimal phase shift of each level (see Figure 3.17).
Next, the aligned shape profiles are converted back to the point cloud in CCS, and the result is compared with the alignment of ICP. Figure 3.18 shows that the proposed multi-resolution alignment method results in smaller root mean square errors (RMSE) than ICP at all four levels. Clearly, the proposed alignment method outperforms ICP in aligning the two-dimensional point cloud.

It is also worth mentioning that the proposed alignment method is comparatively more computationally efficient than the iterative alignment algorithms (e.g., ICP). The wavelet decomposition, alignment, and deviation quantification are not time-consuming which would not take longer than
1 second in total. The potential bottleneck occurs at the alignment optimization procedure. As \( N \) increases, the computational time will gradually increase. In this experiment, it cost 2.35 seconds to obtain the optimal alignment at \( N = 25 \). The trade-off between the performance and computational time is important, however, not a significant issue in this proposed multi-resolution framework.

### 3.4.2 The 2nd layer — a defective layer

The 2nd sample layer, as shown in Figure 14, has a defective area, which is created manually for validating the error quantification capability of the proposed method. The optimal alignment parameters for this layer are \( N = 19 \) and \( L = 3 \). The alignment profiles based on the optimal setting are presented in Figure 3.20.

The point-wise deviations are then calculated based on the alignment at each level. It is observed in Figure 3.21 that the deviation values are small. To reduce the effect of sample size on the result of ANOVA, only a certain number of points are sampled from the population. Again, each
level of alignment is considered as a treatment and each sample point \((\theta)\) is denoted as a block. It is worth mentioning that two sampling strategies can be implemented according to a user-defined inspection requirement.

If a global inspection or alignment consistency check is required, the sample points can be generated uniformly from the population. An example based on 100 regularly sampled points is provided in Table 3.2 and Figure 16. The result of the hypothesis test on the alignment consistency is indicated in the ANOVA table as shown in Table 3.2. Clearly, the p-value is larger than the significant level 0.05, indicating that different levels of alignment are consistent. Based on Eqs 3.7-3.9, the systematic error, alignment error of each level, and random error at each point are calculated and demonstrated in Figure 3.22. The magnitude of each error term is calculated based on the standardized polar coordinates. Therefore, all the errors should be scaled back to their original domain for interpretation in practice.

(a) coefficient profiles    (b) aligned coefficient profiles    (c) aligned shape profiles

Figure 3.20: 3-level multi-resolution alignment results.

Figure 3.21: Deviation profiles.
Table 3.2: Result of ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Adj SS</th>
<th>Adj MS</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment</td>
<td>2</td>
<td>0.00036</td>
<td>0.000180</td>
<td>1.50</td>
<td>0.225</td>
</tr>
<tr>
<td>block</td>
<td>99</td>
<td>1.07970</td>
<td>0.010906</td>
<td>91.14</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>198</td>
<td>0.02369</td>
<td>0.000120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>299</td>
<td>1.10376</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.22: Result of error quantification based on 100 uniformly sampled points.

If a local inspection is required, the points are sampled only from the concerned area defined by users. In this experiment, there are 57 points sampled around the defective area. The error quantification results regarding the systematic, alignment, and random errors are shown in Figure 3.23. One can see that the local systematic error is larger than what is observed in the global inspection. In addition, the alignment error is much smaller than the other two error terms in both global and local inspections.

Note that the errors measured in both global and local inspection are based on the standardized coordinates. Therefore, in practice, all the errors should be scaled up based on the size of the printed sample by multiplying the error with $\max_{\theta \in \Theta} r(\theta) - \min_{\theta \in \Theta} r(\theta)$. In this case, the local systematic error is scaled up as -1.64 unit of the sample size, which is close to the designed...
deviation -1.8 unit as we introduced to the layer. This supports the claim that the proposed multi-resolution alignment method can work effectively even on aligning defective layers.

![Figure 3.23: Result of error quantification based on 57 points around the defective area.](image)

3.4.3 The 3rd layer — a layer with key internal features

The 3rd sample layer is much more complex than the previous two layers. It contains 6 internal features and 1 outer contour shape. The features, as shown in Figure 3.24, can be separated by spectral clustering (Zelnik-Manor and Perona, 2005). In this case, the outer contour shape (cluster 1) is considered as a datum feature for multi-resolution alignment. Therefore, the shape profile of cluster 1 (see Figure 3.25) is decomposed into 3 levels of detail coefficients. The detail coefficients, datum shape profile, and original complex shape are then aligned using $N = 25$ and $L = 3$ (see Figure 3.26). All features are aligned based on the phase shift of the datum.

Based on the aligned shapes, each feature (cluster) is inspected separately by following individual error quantification procedures. The systematic error and alignment error of each feature at different resolution levels are shown in Table 3.3. It can be observed that most clusters, except for
Figure 3.24: Feature segmentation by spectral clustering.

Figure 3.25: Shape profile of the outer contour shape.

Cluster 1, have achieved satisfactory accuracy.

Table 3.3: Error Quantification of Each Feature

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Systematic error</th>
<th>Alignment error (level1)</th>
<th>Alignment error (level2)</th>
<th>Alignment error (level3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>0.0376</td>
<td>-0.0372</td>
<td>0.0727</td>
<td>-0.0355</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>-0.0030</td>
<td>-0.0090</td>
<td>0.0039</td>
<td>0.0051</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>0.0074</td>
<td>0.0001</td>
<td>0.0010</td>
<td>-0.0011</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>-0.0340</td>
<td>0.0232</td>
<td>-0.0470</td>
<td>0.0237</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>0.0081</td>
<td>-0.0001</td>
<td>-0.0011</td>
<td>0.0012</td>
</tr>
<tr>
<td>Cluster 6</td>
<td>0.0014</td>
<td>0.0270</td>
<td>-0.0174</td>
<td>-0.0095</td>
</tr>
<tr>
<td>Cluster 7</td>
<td>-0.0115</td>
<td>-0.0068</td>
<td>0.0046</td>
<td>0.0022</td>
</tr>
</tbody>
</table>
This is one noticeable advantage of the proposed methods that it has the ability in handling complex layers with internal features by conducting deviation quantification in a separate and adaptive fashion. In this work, the outer contour of the complex is considered as the datum feature for alignment. In practice, any feature can be considered as the datum, and the rest features can be aligned based on the alignment of the datum. This capability allows the proposed framework to be involved in broader applications associated with layer-wise inspection of complex geometries especially in biomedical and aerospace industries.

### 3.4.4 A case study of QCD method

To demonstrate the procedure of QCD under the multi-resolution framework, layer 75 to layer 175 from the upper half of the human heart are considered as sample layers to be inspected. A defective area is added to the simulated sample part, where the deviation starts at layer 120 and ends at layer 175 as shown in Figure 3.27.

The systematic error of each layer is calculated under the multi-resolution framework based on 100 uniformly sampled points. Figure 3.28 shows that there is an obvious change in the systematic error.
error distribution. If the change point detection method works effectively, the value of $\hat{l}$ close to 120 can be observed at a post-change model that has a mean value close to the magnitude of the mean of systematic error after layer 120 (equal 1 in this designed defective sample). The ADD calculated based on the post-change model with different mean values $\mu_d$ is demonstrated in Figure 3.29. It is observed that a minimum ADD equal to 1 is achieved at $\mu_d \in [0.3, 1.6]\)
3.5 Conclusion

This work proposed a multi-resolution in-plane alignment and inspection scheme in additive manufacturing. The proposed framework is able to not only distinguish errors generated from different sources but also evaluate the performance and consistency of alignment when varying the level of details. Three example layers of a human heart are used to illustrate the superior performance of the proposed framework over existing methods. Moreover, one of the most important capabilities of the proposed framework is that it can deal with complex layers with internal features by conducting deviation quantification in a separate and adaptive fashion. Regarding on-line applications, the proposed framework can be integrated into an online inspection and monitoring system in additive manufacturing to generate alarms when nonconformities are observed in the current working layer.

References


Chen, G., and Maggioni, M.. ”Multiscale geometric dictionaries for point-cloud data”. In International Conference on Sampling Theory and Applications (SampTA), 4(6), 2011.


4 Optimization of a Concurrent AM Process

Fused filament fabrication, like most layer-wise additive manufacturing processes, is hindered by low production rate and scalability issues. Concurrent fused filament fabrication mitigates these disadvantages by distributing the processing of each layer amongst multiple extruders working in parallel. The objective of this work is to develop a general toolpath allocation and scheduling methodology to achieve this. Breaks in toolpath that are inherently created by slicing software for single-extruder machines are used to form sub-paths, and the assignment of these to available extruders is formulated as a scheduling problem with collision constraints. A formal optimization model is presented, and two novel heuristics are developed to obtain approximate solutions. Three case studies demonstrate the application of these algorithms and compare their relative performance with respect to fabrication time and computational cost. In simulations with three extruders, layer printing times were reduced by as much as 60% compared to single-extruder machines. The proposed heuristics also exceeded the performance of two baseline toolpath scheduling algorithms by as much as 45%. Two key layer characteristics were found to influence heuristic performance, and the advantages and disadvantages of each algorithm are discussed in the context of these characteristics.

4.1 Introduction

Fused Filament Fabrication (FFF) is an additive manufacturing (AM) process based on extrusion, whereby polymer filament is deposited by a moving extruder to form layers (Gibson et al., 2014). As with AM in general, it excels at producing complex geometric forms without the need for part-specific tooling. Compared to other AM processes, it has high material utilization, and it is relatively low-maintenance, safe, inexpensive, and environmentally benign (Gao et al., 2015). As one of the most popular AM techniques, FFF has also been shown to have good potential for industrial applications, such as automobile prototyping and rapid tooling (“3D Printed Shelby Cobra,” n.d.). In addition, reinforcing fibers may be added to the feedstock to increase strength (Hawley et al., 1984), and FFF for carbon-fiber composites based on thermosetting polymers has
been demonstrated (Compton et al., 2014).

The origin of FFF lies in producing small but complex geometries. However, with demand from industries such as aerospace, architecture, and automotive for the low-volume fabrication of large products, there is a strong motivation to scale up the process (Vurpillat, 2016; “BetAbram,” n.d.; Sevenson, 2015; “Erectorbot,” n.d.). Nevertheless, upscaling the FFF process is hindered by limited fabrication speed. Increasing layer thickness can speed up the process, but this sacrifices resolution and increases staisstepping errors (Kim et al., 2008). This limitation forces manufacturers and customers to choose between long cycle times and extremely poor surface finishes.

To improve the scalability of FFF, a significant decrease in workpiece cycle time without degrading its resolution is needed. The most straightforward method adopted in subtractive manufacturing to reduce cycle time is to decompose the task into several subtasks and assign each to different machines working in parallel. In an analogous approach, this research proposes that FFF machines could be equipped with multiple extruders working collaboratively on each layer. In such a concurrent multi-extruder AM system, each extruder would be independently driven, and each would be able to reach the entire workspace in order to maximize flexibility. Existing FFF planning algorithms already divide the toolpath for each layer into several sub-paths, including perimeters for outlining the layer and multiple rasters for interior in-filling (Jin et al., 2015), and these may be used to subdivide the task of printing the layer among the extruders. A key advantage to this approach is that it works with any existing toolpath planning software, effectively decoupling path planning from path scheduling. This allows manufacturers to implement the heuristics developed herein using any software that makes toolpath code available to the user and allows developers of proprietary systems to do the same without the need to adapt or redevelop their toolpath planning software. Further, risk for end users is mitigated by using tried and true toolpath planning software that they already trust to produce quality parts with predictable properties.

There are two major issues that must be resolved before such a concurrent multi-extruder system may be developed. First, since all the extruders share the same workspace, a collision-avoidance strategy must be developed. Second, the makespan (i.e., the time to complete one layer)
of the multi-extruder system is highly dependent on the assignment and sequencing of sub-paths among extruders. Further, these problems are intertwined: Sub-path allocation and scheduling can influence the potential for collision, and collision avoidance will place constraints on sub-path scheduling. To address the two issues, a method for finding the optimal schedule that will prevent collisions while achieving the minimum makespan is of special interest. In addition, this method should be computationally efficient so as not to introduce unnecessary delays prior to printing, between layers, or within a layer.

In this paper, the first issue is solved by a collision check strategy based on the area swept by each extruder over a given time interval. The second issue is addressed by formulating an optimization model for the multi-extruder scheduling problem, with the objective of minimizing the makespan subject to collision constraints. Since the optimization problem is NP-hard, two heuristic algorithms are proposed. The first method features a novel combination of collision avoidance strategies with the Longest Processing Time First (LPT) scheduling algorithm. The second is an entirely novel method, called the Minimum Overlapped Area First (MOA) algorithm, which reduces the likelihood of collision by prioritizing concurrent scheduling of sub-paths having minimal mutual interference. The performance of these heuristics is measured and compared in terms of makespan the computational time through three case studies with layers possessing different characteristics. In these case studies, layer printing time was reduced by as much as 60%

The remainder of this paper is organized as follows: Section 4.2 provides a brief literature review on current multi-extruder FFF applications, collision avoidance in robotic motion planning, and existing parallel scheduling algorithms. In Section 4.3, the optimization model, heuristic methods, and collision avoidance strategies are described in detail. In Section 4.4, three case studies are provided to demonstrate the application of the algorithms and compare their performance. Finally, Section 4.5 discusses the conclusions drawn from the case studies and suggests directions for future research.
4.2 Literature Review

While many FFF machines possess multiple extruders, in the vast majority of cases the extruders are fixed on the same moving platform, meaning they are not capable of independent motion. Such machines are usually designed for multi-material and/or multicolor printing (Pax, 2013; Schumacher et al., 2014). This arrangement can also be used to implement synchronous printing, whereby multiple copies of the same part are created simultaneously on the same platform to speed up the fabrication for batch production (“Cultivate3d,” n.d.). This approach limits part size to a fraction of the build platform and is clearly not a solution for one-off production, so it is of limited use in addressing the fabrication speed problem in large-scale 3D printing. By contrast, concurrent multi-extruder FFF involves extruders which are capable of independent motion and controlled so as to cooperatively complete each layer. One notable example is Project Escher (“Autodesk,” n.d.; Miller, 2016) announced by Autodesk in 2016. Project Escher, however represents but one of the numerous kinematic arrangements that could be chosen for the implementation of concurrent FFF, and its control is simplified considerably by the fact that the extruders may not overlap in the X-Y plane. There is a need for a general treatment of the problem, published in the open academic literature, to drive innovation and advancement of concurrent FFF fabrication technology.

The two key challenges in concurrent FFF are avoiding collision and efficiently allocating tasks between extruders. Fortunately, the literature is rich with studies on collision avoidance and parallel process scheduling, which provides a strong theoretical and technical starting point for formulating the optimization problem and developing computationally efficient heuristics to arrive at a good solution. Collision avoidance is covered extensively in the robotics literature since it is one of the major tasks in multi-agent planning (Arbib and Rossi, 2000; Wu et al., 2002; Klee et al., 2015). The main purpose is to find a collision-free trajectory that achieves the optimization goal in terms of the total distance or travel time, scheduled arrival time, and time delay (Jeevamalar and Ramabal, 2012). The literature on multi-robot collision avoidance mainly focuses on one of three aspects: obstacle avoidance for mobile robots, collision-free trajectory planning for manipulators, or collision avoidance for multiple robot systems. The collision detection techniques
applied in the first two aspects mainly involve sensing and navigating the environment (David et al., 2016; Pandey and Parhi, 2016; Almasri et al., 2016; Chong et al., 2009), graphic mapping (Švestka and Overmars, 1998; Yang et al., 2010; O’Donnell and Lozano-Pérez, 1989; Lee and Lee, 1987), and mathematical modeling and algorithm development (Ata and Myo, 2006; Hasan and Gavrilova, 2010; Daachi et al, 2012). However, there is often more flexibility in the trajectory of robots, whereas an FFF extruder must not vary from its path or speed while extruding. In addition, there is limited work on how task allocation and scheduling among robots dynamically affects the collision check procedure. Babu et al. (2006) invented a cluster tool to allow multiple robots to serve a plurality of processing stations without collision. Each robot’s trajectory forms an envelope and the process is collision free if all envelopes do not overlap. However, the robots have to sense their location and motion in real time in order to find a collision-free path, which is infeasible in the FFF process where extrusion cannot be interrupted mid-path. Nevertheless, the envelope idea provided by Babu’s team inspired the proposed swept-area collision check method proposed in this paper. Jose et al. (2016) used a genetic algorithm for task allocation and an A* algorithm for path planning in a multi-robot system. Unlike the paths that are predetermined by the slicing software in AM, the path generated by the A* algorithm can adapt to different allocation solutions. Since the positions of the robots are fixed when they perform their intended tasks, the collision check is only for the rapid travel motion based on the instantaneous coordinates of the robots in a pre-defined map; a collision is indicated if any two of the robots have the same position coordinates at a particular time. Unlike a multi-robot system, the operation of the FFF process is piecewise continuous, and each sub-path should be considered as an indivisible task. In essence, the nature of the FFF process leads the present work to be formulated as a scheduling problem with collision constraints rather than as a path planning problem, which is dominant in the robotics literature.

The general objective of a parallel scheduling problem is to minimize the makespan \( C_{\text{max}} \), which is defined as the elapsed time between starting and finishing one production unit (Pindeo, 2012). Let \( P_m \parallel |C_{\text{max}} \) denote such a scheduling problem with \( m \) parallel machines. It has been
shown that, in general, this is an NP-hard problem (Garey and Johnson, 1978), and this fact has inspired the development of different heuristics, including branch-and-bound enumeration (Azizoglu and Kirca, 1999; van de Velde, 1993; Belkaid et al., 2012) and local neighborhood search for reassignment (Brucker et al., 1996; Brucker et al., 1997), as well as approximations using list scheduling (LS) (Granham, 1969), bin-packing base (Coffman et al., 1978), and linear programming (Lenstra et al., 1990; Potts, 1985). Among all the heuristics, LS is particularly powerful for solving on-line scheduling problems (Chen et al., 1998). If the jobs are listed in descending order of processing time, then this is known as the longest processing time first (LPT) algorithm (Granham, 1969; Coffman et al., 1978). Compared to the optimal condition (OPT), LPT can achieve the worst-case ratio of $C_{\text{max}}(\text{OPT})/C_{\text{max}}(\text{LPT})=4/3-1/(3m)$ (Coffman et al., 1978). Because of its proven performance, LPT will serve as the starting point for the development of the algorithms presented herein.

Implicit in the previously mentioned scheduling algorithms is the assumption that the machines/processes are independent, yet this is not the case in the concurrent FFF method envisioned herein; collision between extruders working in close proximity is a very real possibility. As for incorporating interdependency constraints into scheduling problems, Miura et al. (2002) proposed a method of assigning and scheduling detection tasks among parallel working cameras in a dynamic environment. Assignments were evaluated by considering the number of tracked persons and the cameras’ attention distribution in the given specific space. Instead of solving the problem of physical collisions between equipment, the method proposed in Miura’s study is intended to solve resource utilization conflicts. Since collisions in the multi-extruder process mainly result from the physical motion of the extruders, the methods proposed in Miura’s work are not well-suited to the present problem. Li et al. (2015) provided a scheduling model for a robotic flow-shop with multiple robots serving parallel machines. However, the overlap indicated in this study is the crossover of robots on the same track. The crossover can be avoided by determining the optimal operation cycle of each robot. Considering the comparatively flexible movement of extruders and uncertain cycle time of printing tasks, this method still cannot address the collision between ex-
truders. Based on this brief review of related literature, an adequate solution to the problem of parallel scheduling for multiple extruders with collision avoidance does not presently exist. This research addresses this problem and makes the following contributions to the literature base:

1. A formal optimization model for minimizing the makespan under collision avoidance constraints is proposed.

2. A novel modification of the LPT algorithm incorporating collision avoidance is presented as one method for finding a near-optimal solution.

3. A novel, and potentially more computationally efficient algorithm, Minimum Overlapped Area First (MOA), is presented.

To the best of the authors’ knowledge, this research is the first attempt to develop an integrated method of the collision checking and parallel scheduling for the FFF process.

4.3 Method

4.3.1 Assumptions

The model and algorithms presented in this section were developed under the following assumptions:

1. Sub-paths are pre-known and derived from single-extruder toolpath discontinuities. Toolpath generated by existing AM slicing software for single-extruder operations commonly includes printing pauses while the extruder rapid travels to a new position within the layer. These pauses define the start and end points of sub-paths which may be assigned to multiple extruders. This approach allows users to utilize existing slicing and toolpath generation tools, even if they are proprietary. Precisely replicating the printing discontinuities that exist in single-extruder operations also means that workpiece quality should be similar, thereby easing the process qualification process.
2. Perimeter sub-paths must precede raster sub-paths. In order to have better surface finish and dimensional accuracy, perimeters are conventionally printed before the infill rasters. This research maintains this convention.

Additionally, several simplifying assumptions have been made so as not to distract from demonstrating the key aspects of the algorithms. Any could be removed by relatively straightforward methods.

1. The kinematic chain to which the extruders are attached is neglected. In reality, the specific mechanism(s) used to manipulate the extruders may introduce additional opportunities for collision. This could be accounted for with a more complex collision checking algorithm that includes all rigid bodies supporting the extruder, though such an algorithm would lack generality since it is tied to a specific printer geometry. In the case of mobile extruder bases (Milkert, 2014) or kinematically redundant manipulators (Chiaverini et al., 2008), the collision check algorithm developed herein is directly applicable since the kinematic chain that supports the extruder can be repositioned and/or reconfigured in real time.

2. Translational velocity of the extruders is constant for a given layer; that is to say that the two dimensional position of the extruder is given by \( x = x_0 + f \Delta t \), where \( f \) is a feed rate vector of constant magnitude and \( \Delta t \) is the collision check time period (discussed in detail in section ). Accelerations due to starting, stopping, and changing the direction of the extruder are neglected since these are typically very high, making them a relatively minor factor in computing extruder trajectories. If accelerations are indeed significant, extruder position may be evaluated in a piecewise fashion using \( x = x_0 + \frac{1}{2} a \Delta t \) for sections of the trajectory subject to acceleration.

3. Rapid travel is neglected. Rapid travel time is typically very short compared to the time that the extruders are actually engaged, and it is assumed that a collision-free rapid travel path may be found using path planning algorithms from the robotics literature, including those mentioned in the literature review. Rapid travel time between all sub-paths may be computed
at the outset and treated as a delay between sub-paths, which is discussed in detail in section 4.3.6.

4.3.2 Brief Description of Three Cases

Two major factors will influence the process makespan: sub-path length variability and the probability of collision. A large makespan may result from an unbalanced schedule due to the large variance in sub-path lengths or delay introduced to prevent collisions. To illustrate these cases, three layers (Figure 4.1) with different features are considered. The first is a simple layer presented for the purpose of demonstrating the algorithms. The second layer consists of sub-paths with large variance in length, while the third one has sub-paths with nearly uniform length. These two cases are used to compare the performance and relative merits of the proposed heuristics with respect to makespan and computational time.

![Figure 4.1: The overview of three different layers](image)

4.3.3 Baseline Methods

Intuition suggests that dividing each layer into contiguous, and inasmuch as possible, equal regions – one for each available extruder – would be a simple but reasonable approach to sub-path allocation. Two versions of this approach are offered as baselines for evaluating the performance of the LPT and MOA heuristics proposed herein. The first simply divides the bounding rectangle of the layer into equal areas without regard to the actual layer geometry. This naïve scheduling (NS) algorithm is simple to automate and extremely computationally efficient. Of course, considering
the actual layer geometry in establishing the contiguous regions should provide a better solution in most circumstances. Assuming that human intuition would be almost as good or possibly better than any image analysis technique for this purpose, the second algorithm relies upon a human to establish regions with simple boundaries. Automating this human segmented scheduling (HSS) algorithm is of course impractical, but beating its performance serves as a formidable challenge for the proposed heuristics. The following procedure is applicable for both baseline methods: (1) Initial boundaries are established. Let $z_i$ represent region $i$ for $i = 1, 2, \ldots, m$, where $m$ is the number of extruders. (2) Sub-paths that lie entirely within a given region are first assigned to that region. Then the total length of sub-paths assigned to each region, denoted as $L_i$, is calculated. (3) The remaining sub-paths, which span multiple regions, are then assigned to the region with smaller $L_i$. The only difference between the NS and HSS methods is in the method of establishing the initial regional boundaries. For the three layers provided in section 4.3.2, the regions generated by the NS algorithm is shown in Figure 4.2 and those of the HSS algorithm are shown in Figure 4.2. After the regions are established, each extruder starts printing the assigned sub-paths from left to right (i.e., the sub-path with the leftmost start point is scheduled first). Whenever a potential collision is detected (the collision check method will be described in section 4.3.7), the candidate sub-path on one extruder will wait until conflicting sub-paths on other extruders have completed.

![Figure 4.2](image.png)

**Figure 4.2:** Region division in the NS method (boundaries represented by dashed lines)
4.3.4 Optimization model

An optimization model can be formulated as a parallel scheduling problem that minimizes the makespan of the process subject to collision-avoidance constraints. A nomenclature for this model is given in Table 4.1. The parallel scheduling problem can be formulated as

\[
\begin{align*}
\text{min} & \quad C_{\text{max}} \\
\text{s.t.} & \quad \sum_{j \in J} p_j x_{ij} \leq C_{\text{max}} \quad \forall i \in M
\end{align*}
\]  

\[S \equiv \{(j, k) : \chi_{ij} + \chi_{ik} > 1, \quad \forall i \in M\}\]

\[D \equiv \{(j, k) : x_{ij} + x_{ik} \leq 1, \quad \forall i \in M\}\]

\[ST_k = \sum_j a_{jk} p_j ET_k = \sum_j a_{jk} p_j + p_k \quad \text{for} \quad (j, k) \in S\]

\[C \equiv \{(j, k) : (j, k) \in D | ST_k \in [ST_j, ET_j] \quad \text{or} \quad ST_j \in [ST_k, ET_k]\}\]

\[CST_{jk} = \max\{ST_j, ST_k\} \quad \text{for} \quad (j, k) \in C\]

\[CET_{jk} = \min\{ET_j, ET_k\} \quad \text{for} \quad (j, k) \in C\]

\[\sum_g \varphi_{jkg} = 0 \quad \text{for} \quad G_{jk} = \frac{CET_{jk} - CST_{jk}}{dt} \quad \text{and} \quad (j, k) \in C\]
\[
\sum_{i \in M} x_{ij} = 1 \quad \text{for} \quad \forall j \in J
\]  
\[
a_{jk} + a_{kj} \leq 1 \quad \text{for} \quad j, k \in S
\]

where \(C_{\text{max}}\) is the makespan, which should be longer than the completion time of any individual extruder. \(p_j\) is the processing time for sub-path \(j\). Two collections of decision variables are defined for this problem: \(\chi_{ij}\) and \(a_{jk}\). Note that for any \(a_{jk}\), sub-path \(j\) does not necessarily immediately precede sub-path \(k\).

Collision checking is accomplished by considering the area swept out by each extruder over a given time period. In order to refine the sub-path set for collision checking, we further define two sets: \(S\) and \(D\), where set \(S\) includes pairs of sub-paths assigned to the same extruder and set \(D\) contains those on different extruders. The start time \(ST_k\) of sub-path \(k\) equals the total processing time of sub-paths ahead of \(k\) on the same extruder, and its end time \(ET_k\) can be obtained by adding the processing time of sub-path \(k\) to \(ST_k\). If the processing period for any pair of sub-paths \((j, k)\) in set \(D\) has an overlapped period \([CST_{jk}, CET_{jk}]\), a collision check is performed by dividing the time period into \(G_{jk}\) time slots, where \(G_{jk}\) is determined according to a predefined time interval \(dt\). If a potential collision between sub-path \(j\) and \(k\) is identified, the collision sign \(\varphi_{jk\ell}\) for any time slot \(\ell\) in this time period equals 1. The detailed collision check process is illustrated in section 4.3.7. Additionally, perimeter sub-paths should be printed prior to infill raster sub-paths. Thus, the minimum makespan for printing perimeter sub-paths is first calculated by including all the perimeter sub-paths in the set \(J\). Let \(C_P\) denote the completion time of all perimeter sub-paths, and let the set \(J\) equal to a set \(R\) that includes all the raster sub-paths and modify equation (2) as \(C_P + \sum_{j \in R} p_j x_{ij}\) for \(\forall i \in M\).

### 4.3.5 Heuristic Algorithms

To solve the NP-hard parallel scheduling problem, two collision-free heuristic scheduling algorithms are developed that make the problem tractable. The first one combines the longest processing time first (LPT) with collision avoidance constraints; the second algorithm, minimum
Table 4.1: Symbols Used in the Optimization Model

<table>
<thead>
<tr>
<th>Set</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J$</td>
<td>A given set of sub-paths. $j, k \in J$</td>
</tr>
<tr>
<td>$M$</td>
<td>A given set of available extruders. $i \in M$</td>
</tr>
<tr>
<td>$S$</td>
<td>Include pairs of sub-paths $(j, k)$ assigned to the same extruder. It will change with the decision variable.</td>
</tr>
<tr>
<td>$D$</td>
<td>Include pairs of sub-paths $(j, k)$ assigned to the different extruder. It will change with the decision variable.</td>
</tr>
<tr>
<td>$C$</td>
<td>Include pairs of sub-paths $(j, k)$ that need collision check. It requires $(j, k) \in D$ and the processing periods of sub-path $j$ and sub-path $k$ overlap in the time domain.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{ij}, x_{ik}$</td>
<td>$x_{ij} = {1$, if sub-path $j$ is assigned to extruder $i$; $0$, otherwise$}, j \in J, i \in M$</td>
</tr>
<tr>
<td></td>
<td>$x_{ik} = {1$, if sub-path $k$ is assigned to extruder $i$; $0$, otherwise$}, k \in J, i \in M$</td>
</tr>
<tr>
<td>$a_{jk}$</td>
<td>$a_{jk} = {1$, if sub-path $j$ is assigned before sub-path $k$; $0$, otherwise$}$. Note that for any $a_{jk}$, sub-path $j$ does not necessarily immediately precede sub-path $k$.</td>
</tr>
<tr>
<td>$ST_k$</td>
<td>The start time sub-path $k$, which is equal to the total processing time of sub-paths $j$ ahead of $k$ on the same extruder. $(j, k) \in S$</td>
</tr>
<tr>
<td>$ET_k$</td>
<td>The end time sub-path $k$, which can be obtained by adding processing time of sub-path $k$ to the start time sub-path $k$.</td>
</tr>
<tr>
<td>$CST_{jk}$</td>
<td>The start time of overlapped period for $(j, k) \in C$</td>
</tr>
<tr>
<td>$CET_{jk}$</td>
<td>The end time of overlapped period for $(j, k) \in C$</td>
</tr>
<tr>
<td>$G_{jk}$</td>
<td>The number of time slot for collision check, which is obtained by divide the length of the overlapped period with a given time interval $dt$</td>
</tr>
<tr>
<td>$\varphi_{jkg}$</td>
<td>The collision sign for any pair of sub-paths $(j, k) \in C$ during any time slot $g$, where $g = 1, 2, \ldots G_{jk}$; $\varphi_{jkg} = {1$, if a potential collision is identified between sub-path $j$ and $k$ during time slot $g$; $0$, otherwise$}$.</td>
</tr>
<tr>
<td>$C_{max}$</td>
<td>Makespan of the printing given set of sub-paths, which should be longer than the completion time of any individual extruder.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_j, p_k$</td>
<td>Process time of sub-path $j, k$, where $j, k \in J$</td>
</tr>
<tr>
<td>$dt$</td>
<td>A given time period for determining the size of the time slot for collision check</td>
</tr>
</tbody>
</table>

overlapped area first (MOA), is a novel method for assessing the potential for collision between sub-paths and attempting to simultaneously schedule sub-paths that have minimal potential for collision. Both of the heuristic algorithms are generally applicable as long as all the sub-paths are pre-known.

For better surface finish and dimensional accuracy, all perimeter sub-paths should be completed before starting the infill raster sub-paths. Thus, for each layer, the scheduling of perimeters and
rasters can be considered as two separate, sequential problems.

**Longest Processing Time First (LPT)** In solving a regular parallel scheduling problem $P_m || C_{max}$, the conventional LPT algorithm assigns the $m$ longest jobs to $m$ machines at $t = 0$. After that, the longest job among those not yet processed is assigned to each machine as the machine becomes free (Pindeo, 2012). In this research, the LPT algorithm was modified by adding a collision checking step (details provided in Section 4.3.7). Figure 4.4 shows a flow chart of the algorithm, described as follows: The modified LPT algorithm starts by assigning the longest sub-path to the first extruder. Then it searches the remaining sub-paths in descending order of processing time until a sub-path is found that will not cause a collision with those currently being processed. This newly found sub-path is assigned to a free extruder, and the process is repeated until all extruders have been assigned a sub-path. Assuming that the number of sub-paths exceeds the number of extruders, the earliest time at which an extruder will complete its assigned sub-path and become free is computed. The algorithm then searches for the longest unassigned path that will not generate a collision if scheduled at that time. This process repeats until no other sub-paths can be scheduled for parallel processing, and a single extruder completes the remaining sub-paths for the layer. Since perimeters must be completed before rasters, the modified LPT algorithm is applied twice: first for perimeters, then for rasters. It is worth pointing out that the modified LPT algorithm requires repeated collision checks, which can become computationally expensive for some layer geometries. This weakness motivated the development of an alternative algorithm.

**Minimum Overlapped Area First (MOA)** In general, the potential for collision between any given pair of sub-paths is related to the intersection of the areas swept out by the extruders while completing those sub-paths. A smaller overlap indicates a lower probability of collision during concurrent processing, and no overlap means that no collision will occur. The general principle behind MOA is to concurrently schedule sub-paths with minimal overlap. This algorithm can reduce the number of collision checks in some layer geometries, and lower the computational cost accordingly. A flowchart for the MOA algorithm is shown in Figure 4.5. Initially, the total
overlapped area between each sub-path and all other sub-paths is computed, and the one with the highest total overlapped area is scheduled on the first extruder. This has the effect of scheduling the most troublesome sub-path as early as possible, which was found to improve algorithm performance. Subsequent sub-paths are assigned by first ranking them in order of overlapped area with the sub-paths currently in progress, then searching these from the lowest to the highest for
a sub-path that will not produce a collision. The calculation of overlapped areas for all possible sub-path pairs is accomplished in a preprocessing step that precedes the implementation of the MOA algorithm. Each pairwise overlap calculation involves three steps: First, the convex hull of each sub-path is constructed (Figure 4.6a). Then each convex hull is offset by the size of extruder (Figure 4.6b). Lastly, the overlapped area for each possible pair of offset convex hulls is calculated (Figure 4.6c). It is worth pointing out that these calculations need not be repeated in executing the algorithm. The total overlapped area (with all other sub-paths) for use in the initial assignment is derived from these results. As with the LPT algorithm, perimeter sub-paths need to be completed before rastering the infill can begin. The MOA heuristic was found to perform poorly when scheduling perimeters due to a very high incidence of large overlapping areas of the convex hulls. As a result, perimeter scheduling is always accomplished via the LPT algorithm, even when MOA is used for the rasters.

4.3.6 Adaptive Algorithm Modification

In both the LPT and MOA algorithms, it is quite possible that as a given extruder becomes free, all remaining unscheduled sub-paths will cause collision if immediately scheduled. In other words, the collision sign of $Res[j]$ in Figure 4.4 (LPT) or MOA[$j$] Figure 4.5 (MOA) for any $j \in Res$ will be 1. Figure 4.7 shows an example: Extruder 2 finishes sub-path A and is ready to accept a new sub-path, but scheduling any of the sub-paths in the set of those that remain, $Res = \{1, 2, 3\}$, would result in one or more collisions with B and C, already in process. In the basic algorithms described thus far, Extruder 2 would be idle until the completion of C, at which time the algorithm would evaluate whether any of the remaining sub-paths could be scheduled on Extruders 2 and/or 3. If not, none of the remaining sub-paths could be scheduled until the completion of B, leaving both 2 and 3 idle. In practice, it was found that a number of incompatible sub-paths were frequently scheduled sequentially on a single extruder toward the completion of the layer. This left the remaining extruders idle for an extended period of time, partially defeating the purpose of having multiple extruders.
This weakness can often be rectified by shifting the scheduling of a candidate sub-path in time so as to avoid collision. This is done by defining an adaptive delay interval, $\Delta T$. Instead of waiting for the completion of a sub-path, the algorithms evaluate whether each remaining sub-path can be scheduled at time $n\Delta T$ for $n = 1, 2, 3, \ldots$ without causing collision, and the sub-path requiring minimum delay is scheduled. Returning to the example in Figure 4.7, each sub-path in
Figure 4.6: Computing overlapped area for two sub-paths

$Res = \{1, 2, 3\}$ is tested at $\Delta T, 2\Delta T, 3\Delta T$, etc. to determine the earliest time at which it can be scheduled without collision. It is found that 3 can be scheduled after a delay of $2\Delta T$, while sub-paths 1 and 2 require delays of $7\Delta T$ and $4\Delta T$, respectively. Adaptive delay significantly increases the number of collision checks required for each algorithm. Thus, the magnitude of $\Delta T$ represents a tradeoff between minimizing makespan and the computational load of the algorithms.

Figure 4.7: Example of finding the minimum delay

delay = 0

Extruder 1

Extruder 2

Extruder 3

delay

max delay = $7\Delta T$
4.3.7 Collision Check Algorithm

Collision checking is accomplished by considering the area swept out by each extruder over a given time period \( \Delta t \). In general, this area is a polygon whose specific geometric features are determined by the size and trajectory of the extruder. A non-zero intersection between these areas for any two sub-paths is taken to mean that a collision exists between those sub-paths during time period \( \Delta t \). Of course, this is not precisely true – in actuality, a non-zero intersection indicates the existence of potential collision – but if \( \Delta t \) is small, the chances of a false positive is also small. There is a tradeoff, though, in that a smaller \( \Delta t \) increases the number of collision check operations for a given sub-path, which in turn increases computational cost. Further, finding the intersection of general polygons, especially concave polygons which are frequently encountered in this context, can be computationally expensive. For this reason, the area is extended to be the bounding rectangle (Figure 4.8) of the actual swept area for the purpose of computing collisions. Again, this decision represents an attempt to balance computational efficiency versus false positives.

Figure 4.8: False positive example for collision detection. The shadowed area is the true area swept by the extruders.

The procedure for computing the swept area includes three steps:

**Step 1:** Identify the inspection period. Let \( st_{j_1j_2} \) denote the start time of the inspection time period for any pair of parallel processed sub-paths \( j_1 \) and \( j_2 \), which equals the latest start time of the two sub-paths. Let \( ed_{j_1j_2} \) denote the end time, which is the earliest completion time between sub-paths \( j_1 \) and \( j_2 \). An example is shown in Figure 4.9: Extruder 2 is free and ready to accept sub-path A while the other extruders are processing sub-paths B and C. The inspection time period of sub-paths A and B is \([st_{AB}, ed_{AB}]\), and that of sub-paths A and C is \([st_{AC}, ed_{AC}]\).
Figure 4.9: Example of identifying inspection time period

**Step 2:** Divide the inspection time period into several time intervals and find the corresponding path segments that belong to each interval. Letting $\Delta t$ be the length of each time interval, the inspection time period $[st, ed]$ can be divided into

$$T_1 : [st, st + \Delta t] \quad T_2 : [st + \Delta t, st + 2\Delta t] \ldots$$

$$T_i : [st + (i-1)\Delta t, st + i\Delta t] \ldots T_N : [st + (N-1)\Delta t, ed]$$

(4.12)

where $N = \lceil \frac{ed_{j_{1j2}} - st_{j_{1j2}}}{\Delta t} \rceil$. Given sub-path $j$ with $G$ line segments in total and a constant printing speed $\nu$, the corresponding path segment that belongs to any $T_i$ can be found using the procedure described by the following pseudo code. In addition, an example is given in Figure 4.10.
**Algorithm 1 Collision Check Algorithm**

% Let \( D_g \) denotes the length of segment \( g \), \((SX_g, SY_g)\) denotes its start point, and \((EX_g, EY_g)\) denotes its end point. \( g = 1, 2, \ldots G \).
% Assume the start printing time of sub-path \( j \) is \( ST_j \).
% Sweeping area of \( T_1 \):
% The printed length of line segment before \( st \) is \( \Delta DS = (st - ST_j) \nu \)

\( g = 1; \)

while \( \Delta DS > D_g \) do
  \[
  \Delta DS = \Delta DS - D_g;
  \]
  \( g + 1 \to g; \)

% When iteration ends, \( g = gs \). Calculate the start point \( P \) of sweeping area
\[
X_P = SX_g + \Delta DS \frac{EX_g - SX_g}{D_g};
\]
\[
Y_P = SY_g + \Delta DS \frac{EY_g - SY_g}{D_g};
\]
\[
D_g = \text{dist}[(X_P, Y_P), (EX_g, EY_g)];
\]
% the length of segment to be covered in \( T_1 \) is \( \Delta D = \Delta t \cdot \nu \)

for \( i = 1 : N \) do
  while \( \Delta D > D_g \) do
    \[
    \Delta D = \Delta D - D_g;
    \]
    Add segment \( g \) to set \( T_1 \);
    \( g + 1 \to g; \)

  % calculate the end point \( Q \)
  \[
  X_Q = SX_g + \Delta D \frac{EX_g - SX_g}{D_g};
  \]
  \[
  Y_Q = SY_g + \Delta D \frac{EY_g - SY_g}{D_g};
  \]
  \[
  D_g = \text{dist}[(X_Q, Y_Q), (EX_g, EY_g)];
  \]
  % remaining length of segment \( g \)
  % \( Q \) becomes the new start point
  \[
  (X_Q, Y_Q) \to (X_P, Y_P);
  \]

In the example, \( P \) is the start point of line segment in interval \( T_1 \), which is on segment \( g : S_g \to E_g \). The remaining length is \( D_g = \text{dist}(P, E_g) \). After the first iteration, \( \Delta D = \Delta t \cdot \nu - D_g - D_{g+1} \) and \( g = g + 2 \). Thus, end point \( Q \) is on segment \( g + 2 : S_{g+2} \to E_{g+2} \). The coordinates of \( Q \) would be \( (X_Q = SX_{g+2} + \Delta D \frac{EX_{g+2} - SX_{g+2}}{D_{g+2}}, Y_Q = BY_{g+2} + \Delta D \frac{EY_{g+2} - SY_{g+2}}{D_{g+2}}) \). Line segment in \( T_1 \) should include \( P \to E_g, S_{g+1} \to E_{g+1} \) and \( S_{g+2} \to Q \).

**Step 3:** Build the bounding rectangle to form the extended sweeping area. In step 2, all the line segments that belong to any interval \( T_i \) are found. The boundary of the bounding rectangle should be the extreme points of the line segments, offset outward by half of the extruder dimension. For example, Figure 4.11. shows the continued example. The extreme points are indicated by red dots. If the extruder dimension is \( 2r \), then the upper right boundary of the bounding rectangle would be
Figure 4.10: Example of identifying inspection time period

\( X_p + r \) and \( Y_p + r \), respectively. The collision inspection for each pair of sub-paths tests whether the bounding rectangles for time interval \( T_i \) intersect. Intersection detection is similar to the third preprocessing step for calculating the overlapped area in the MOA method.

Figure 4.11: Finding the bounding rectangle to form the extended sweeping area

4.4 Result and Discussion

The proposed heuristics are implemented in Matlab on the three layers described in Section 4.3.2 and their performance is compared to the two baseline algorithms. The first case (Section 4.4.1) is used for demonstrating the LPT and MOA algorithms as well as the methods for comparing their performance. The second and third cases (Sections 4.4.2 and 4.4.3) are presented to com-
pare the relative merits of the proposed heuristics in handling layers with divergent (Case 2) or uniform (Case 3) lengths of sub-paths. In these studies, the sub-paths (including both perimeters and rasters) are generated by the same slicer. The printing speed is 2 units/sec, and the time interval $\Delta t$ for collision checking is 0.1 sec. For comparing the performance of the algorithms, the computational time and makespan (time to print one layer) are of primary interest. In particular, the computational time is calculated based on a desktop computer with an Intel core i7 CPU. In addition, the total delay introduced into the process (as per Section 4.3.6) is also presented, as it adversely affects the computational time.

It is worth pointing out that since MOA is applied only to raster sub-paths, only the raster sub-paths are considered in the second and third cases for performance comparison. For a conceptual investigation, different ratios of extruder size $E$ to layer size $W$ are also considered in the two cases when comparing the performance of the scheduling algorithms. As this ratio increases, finding a collision-free scheduling solution becomes more difficult. For illustration purposes, the size of extruder is changed while the layer size, defined as the layer’s minimum side length $W$, is held constant. However, in practice, with an extruder of fixed size, the ratio varies as the layer size changes.

### 4.4.1 Case 1

The first layer is a 20 unit $\times$ 20 unit circle that includes some concave features. The slicer identified 7 perimeter paths and divided the infill raster into 4 sub-paths (i.e., 11 paths in total as shown in Figure 4.12). The width of the extruder’s bounding rectangle $E$ is 2 units.

Under the given configuration, the LPT and MOA method was applied for up to six extruders, with the resulting makespans and computational times presented Table 4.2. While makespan is reduced by concurrent processing, the improvement asymptotically reaches a limit as the number of extruders is increased. Meanwhile, computational time tends to increase with the number of extruders with a few exceptions. Without declaring an optimal number, which would involve some situationally dependent weighting between the two parameters, three extruders appears to
Figure 4.12: Sub-paths of the layer (where axis values indicate the position of layer on the workspace)

be a reasonable compromise for case 1. Similar analyses for cases 2 and 3 indicate three or four extruders. For the purposes of this paper, a three-extruder configuration is adopted throughout to facilitate comparison between algorithms and cases.

Table 4.2: Performance of LPT and MOA for Case 1 vs Number of Extruders

<table>
<thead>
<tr>
<th>Number of extruders</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LPT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Makespan (s)</td>
<td>318</td>
<td>168</td>
<td>135</td>
<td>131</td>
<td>131</td>
<td>131</td>
</tr>
<tr>
<td>MOA</td>
<td>318</td>
<td>176</td>
<td>135</td>
<td>131</td>
<td>131</td>
<td>131</td>
</tr>
<tr>
<td>Computational Time (s)</td>
<td>0</td>
<td>0.325</td>
<td>2.61</td>
<td>6.81</td>
<td>6.74</td>
<td>6.99</td>
</tr>
</tbody>
</table>

Figure 4.13 shows the path allocations and scheduling results for the three extruders. For bet-
Figure 4.13: Graphic illustration of the scheduling result from the LPT heuristic (a) The toolpath path is shown in time (vertical axis) indicating no intersection between extruder trajectories. (b) The relative size of the extruder’s bounding rectangle to the layer is provided.

ter visualization, both two-dimensional and three-dimensional plots are provided to illustrate the path distribution and progression of each extruder over time. Clearly, the collision-free LPT algorithm effectively avoids collisions between the three extruders in this concurrent AM environment. Similar results for the MOA and both baseline methods may be obtained similarly.

Figure 4.14 shows the scheduling results of all four algorithms visualized in the time domain in order to explicitly show their differences (negative signs indicate the paths were reversed to avoid collision). As previously mentioned, all perimeters should be completed before rastering starts, and MOA is applicable only to raster sub-paths, so differences between the LPT and MOA exit only for sub-paths 8–11. The detailed results of scheduling performance for the four algorithms are shown in Table 4.3.

| Algorithm Performance Comparison among LPT, MOA, NS and HSS for the Simple Layer |
|--------------------------------------|------|-------|------|------|
|                                      | LPT  | MOA  | NS   | HSS  |
| Cmax (sec)                           | 135  | 135  | 245  | 266  |
| Total delay in perimeter (sec)       | 0.3  | 0.3  | 54.0 | 64.8 |
| Total delay in raster (sec)          | 5.2  | 1.2  | 73.2 | 83.2 |
| Computational time (sec)             | 2.90 | 1.47 | 0.56 | 0.38 |

Algorithm performance may be evaluated by comparing to both the single-extruder makespan
Figure 4.14: Result of LPT, MOA, NS and HSS for 3-extruder scheduling. Negative signs indicates that the path was reversed to avoid collision. Black indicates delay introduced to avoid collision.
and the theoretical lower bound of the multi-extruder makespan (single-extruder makespan divided by the number of extruders). For case 1, the single-extruder makespan is 318 s, and the theoretical lower bound is 106 s. Thus, the results represent a 56% improvement over the single-extruder makespan and are 24% slower than the theoretical minimum.

Total delay, defined as the cumulative length of all delay periods on all extruders, is reported in Table 4.3 as measure of mutual interference between extruders. Delay results from the algorithms’ search for a future time at which to schedule an unassigned sub-path. This search involves numerous collision checks, potentially causing a significant increase in computational time. Total delay time assists in differentiating between long computational times are the result of a large layer versus those caused by an overly crowded environment. More delay time in a layer does not necessarily imply an increase makespan, which may seem counterintuitive at first glance. A simple example can be found in Figure 4.14 (a) and (b): although the two schedules have different total delay times, their makespan is identical since the delay time for extruder 1 is identical.

### 4.4.2 Case 2

Case 2 (Figure 4.15) consists of 17 raster sub-paths with high variability in length. The minimum side size $W$ of the layer is 25 units. The width of the extruder’s bounding rectangle $E$ is incrementally increased from 4 units to 12 units to evaluate the algorithms as collisions become more likely. The performance measures of the algorithms are presented in Figure 4.16. By means of comparison, the single-extruder makespan for this layer is 265 seconds.

In this case, the variance in length is significant, so when the ratio of extruder size to layer size is small, LPT performs better than MOA in minimizing the makespan. In other words, the dominating factor that influences the makespan in this case is the sub-path lengths. The longer makespan of MOA is attributed to the unbalanced schedule where some long sub-paths are scheduled at the end of the process. However, as the $E/W$ ratio increases, LPT loses its advantage as additional opportunities for collision introduce long delays and long computational time becomes the dominating factor. As a result, the performance of MOA becomes much better than LPT, not
Figure 4.15: The layer overview of Case 2.

only in minimizing the makespan, but also in reducing the computational time. Both LPT and MOA outperform the baseline methods with respect to makespan under all conditions.

Figure 4.16: Performance Comparison among LPT, MOA, NS, HSS for Case 2

The single-extruder makespan for case 2 is 264 s and the theoretical minimum is 88 s. The
results show a decrease from the single-extruder makespan of between 53% and 67%, and are 0% to 41% slower than the theoretical minimum.

4.4.3 Case 3

The third case study considers the layer shown in Figure 4.17. Sixteen square blocks lead to a small variance in sub-path lengths. The minimum side size \( W \) of the layer is 60 units. In order to study the performance of the algorithms as the probability of collisions increases, the width of the extruder is incrementally increased from 10 units to 24 units. The performance of the two heuristics is summarized in Figure 4.18. By means of comparison, the single-extruder makespan for this layer is 2358 seconds.

![Figure 4.17: The layer overview of Case 3](image)

In this case, all the sub-paths have almost the same length, with variance less than 1 unit. Figure 18 shows that LPT loses its advantage and becomes inferior as the ratio \( E/W \) increases, while on the other hand, the performance of MOA is quite stable with a makespan of 884 seconds. In addition, the computational times of both LPT and MOA increase as the ratio increases, but the increase in LPT is much more significant than that of MOA.

It is notable that even though the baseline methods underperform LPT and MOA in minimizing the makespan as the ratio \( E/W \) increases, the difference is not as significant as it was in the first two cases. This is mainly due to the equal length of the sub-paths. In this case, the maximum possible
delay is the fabrication time of one square block (sub-path). As the delay of LPT increases, the makespan will approach that of the baseline method.

The single-extruder makespan for case 3 is 2358 s and the theoretical minimum is 786 s. The results show a decrease from the single-extruder makespan of 63%, and are 12% slower than the theoretical minimum.

![Figure 4.18: Performance Comparison among LPT, MOA, NS and HSS for Case 3](image)

**4.4.4 Discussion**

The case studies show that the proposed LPT and MOA algorithms reduce fabrication time significantly as compared to the two baseline scheduling approaches, though this performance advantage comes with a computational cost. Given that fabrication time usually far exceeds digital preprocessing time, and that adding computational resources is usually cheaper than adding fabrication
resources, this represents an attractive tradeoff.

The case studies also suggest that two key layer characteristics determine which heuristic will perform better: layer size relative to extruder size (E/W ratio) and the variability of sub-path processing times. Case 2 indicates that the LPT algorithm tends to perform better when the layer is large relative to the extruder size and there is a wide variety of sub-path processing times. These circumstances minimize the number of potential collisions and increase the probability of finding a compatible sub-path that may be scheduled immediately, thus achieving the proven efficiency of the LPT algorithm with reasonable computational cost. If the layer is small relative to the extruder, as in Cases 2 and 3 where the E/W ratio is large, or if the variability in sub-path lengths is low, as in Case 3, extruder utilization and computing time both suffer under LPT. This is due to the fact that LPT does not consider the probability of collision between two paths, and it can expend considerable computing resources in a brute-force effort to schedule the next-largest sub-path, even if it has extensive interference with those already scheduled. The results of Cases 2 and 3 further indicate that the MOA algorithm is more effective and computationally efficient when the potential for collision is high due to a small layer or highly uniform sub-path lengths. This is attributed to the fact that the MOA algorithm gives preference to scheduling concurrent paths with low probability of collision, thereby avoiding excessive searching for minimum delays. But since the MOA algorithm does not consider the processing time for each sub-path, it can postpone long sub-paths toward the end. With fewer unscheduled sub-paths to pick from, finding one that can be concurrently scheduled with this long sub-path can be difficult or impossible, thus increasing makespan and/or computational load.

Based on these results, a hybrid heuristic was developed with the hope of achieving superior performance by combining the best attributes of both algorithms. This algorithm defaults to LPT, thereby assigning higher priority to longer sub-paths. When the algorithm is unable to immediately schedule the next-longest sub-path, it temporarily switches to MOA instead of spending excessive time searching for an acceptable delay. Unfortunately, the hybrid algorithm did not show a performance advantage in the above case studies, and in some instances led to inferior performance.
Both heuristics potentially add time to the digital preproduction process, and the practicality of this should be addressed. The processing times reported herein are based on executing the algorithms in Matlab, and they could be significantly reduced by optimizing the code and switching to a compiled language. It is also possible to parallelize the scheduling and production activities by computing the tool path schedule for layer \(i\) while printing layer \((i-1)\).

4.5 Conclusion

The objective of this research was to create a general toolpath allocation and scheduling methodology to facilitate parallel scheduling of independently-operating extruders for concurrent FFF. This challenge was formulated as a scheduling problem with collision constraints, and an optimization model was developed to minimize fabrication time. Two heuristics were proposed to find approximate but efficient solutions to this otherwise NP-hard problem. One heuristic is a novel modification of the well-established LPT scheduling algorithm, adapted to handle collision prevention. The other (MOA) is a novel algorithm that inherently attempts to concurrently schedule sub-paths with low mutual interference. The application, effectiveness, and computational cost of these heuristics have been demonstrated for three AM layers, and the relative merits of each has been examined and compared with two baseline scheduling approaches.

An advantage of approaching the allocation of toolpath as a scheduling problem is that it decouples toolpath allocation from toolpath planning. The heuristics developed in this paper work with any single-extruder toolpath planning algorithm. This allows AM equipment and software manufacturers to integrate the heuristics directly with their own proprietary toolpath generation code without modification. End users who can access AM G-code (as with the popular Marlin firmware, for example) can implement the heuristics even if they are using closed-source toolpath planning software. These heuristics will also not need to be modified or updated to accommodate new toolpath planning algorithms as they are developed.

Concurrent extruder scheduling can significantly decrease the cycle time for FFF and mitigate its scalability issues. Fabrication times for the layers considered in this paper were reduced by
37% to 66% using three extruders. For large-scale 3D printing, where there is ample space to accommodate multiple extruders, significant reduction in fabrication time is achievable. At present, fabricating large parts with a single extruder implies that layer thickness must scale with part size to keep printing time reasonable. The resulting surface finish on a workpiece the size of an automobile, for example, is so poor that the required post-process finishing operations offset most, and sometimes all, of the benefits of AM. In concurrent FFF, however, the number of extruders can scale with part size, thereby allowing reasonable cycle times with far better layer resolution. Combined with current research on FFF for fiber-reinforced and thermosetting materials, AM of large carbon fiber reinforced polymer parts such as aircraft components or boat hulls is conceivable.

References


5 Summary

Additive manufacturing (AM) has attracted increasing attention and interest of different industrial sectors for its ability of producing complex geometries outside of traditional production environments. However, there are still some technical challenges that need to be solved to reach AM’s full potential. In this dissertation, three important and challenging issues in terms of geometric quality inspection and process efficiency are discussed in detail. Different methodologies applying knowledge of computer vision, statistics, and optimization modeling are proposed in this dissertation to fill several research gaps in offline inspection, online inspection, and concurrent 3D printing.

Among the studies with regard to geometric inspection, this dissertation is the first attempt to develop a fast, feature-based geometric offline inspection technique that does not require alignment before inspection and can be implemented by users without specialized training in inspection according to geometric dimensioning and tolerancing conventions. In addition, this dissertation fills a research gap of state-of-the-art studies regarding multi-resolution analysis of point cloud data for automatically quantifying different types of error inherent in manufacturing, measuring, and part alignment under an online inspection paradigm. Furthermore, this dissertation is among the pioneer studies that first propose and investigate the multi-extruder concurrent AM processes to improve fabrication efficiency while maintaining geometric quality. The contributions of this dissertation have been revealed from both theoretical and practical perspectives based on simulation and physical experiments.

In the future, the methods and models proposed in this dissertation can be integrated into physical AM systems to inspect, monitor, and optimize the process based on real-time feedback. Although the proposed methodologies are discussed under AM application in this dissertation, they can be easily accommodated to other manufacturing processes which are faced with similar technical challenges of the geometric inspection and process efficiency improvement.