An Artificial-intelligence/Statistics Solution to Quantify Material Distortion for Thermal Compensation in Additive Manufacturing

Chao Wang\textsuperscript{a}, Shaofan Li\textsuperscript{a,∗}, Danielle Zeng\textsuperscript{b}, Xinhai Zhu\textsuperscript{c}

\textsuperscript{a}Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720, USA
\textsuperscript{b}Ford Motor Co., Detroit, MI 48226, USA
\textsuperscript{c}ANSYS Livermore Software Technology (ANSYS-LST), Livermore, CA94551, USA

Abstract

In this paper, we introduce a probabilistic statistics solution or artificial intelligence (AI) approach to identify and quantify permanent (non-zero strain) continuum/material deformation only based on the scanned material data in the spatial configuration and the shape of the initial design configuration or the material configuration. The challenge of this problem is that we only know the scanned material data in the spatial configuration and the shape of the design configuration of three-dimensional (3D) printed products, whereas for a specific scanned material point we do not know its corresponding material coordinates in the initial or designed referential configuration, provided that we do not know the detailed information on actual physical deformation process. Different from physics-based modeling, the method developed here is a data-driven artificial intelligence method, which solves the problem with incomplete deformation data or with missing information of actual physical deformation process. We coined the method is an AI-based \textit{material deformation finding} algorithm.

This method has practical significance and important applications in finding and designing thermal compensation configuration of a 3D printed product in additive manufacturing, which is at the heart of the cutting edge 3D printing technology. In this paper, we demonstrate that the proposed AI continuum/material deformation finding approach can accurately find permanent thermal deformation configuration for a complex 3D printed structure component, and hence to identify the thermal compensation design configuration in order to minimizing the impact of temperature fluctuations on 3D printed structure components that are sensitive to changes of temperature.

\textbf{Keywords:} Additive manufacture, Artificial intelligence, Coherent point drift (CPD) method, Data-driven modeling, Material deformation finding algorithm, Maximum likelihood estimate, Maximum a Posteriori estimate, 3D printing

1. Introduction

Due to some of its great technological advantages, Additive manufacturing (AM) has been an international hot-spot in the past several years, such as unparalleled design freedom and short

\textsuperscript{∗}Corresponding Author, Email: shaofan@berkeley.edu
Figure 1: An example of thermal distorted 3D printed product: Distortion on a 3D printed Inconel component heated only from below (left) versus the same Inconel component that was made when heated from above and below (right). (From https://www.pinterest.com/pin/411305378465700906/)

lead time [8]. The unique edge of three-dimensional (3D) printing technology lies in its ability to directly fabricate complex shapes and eliminate the waste that are associated with the conventional manufacturing [4]. In 3D printing manufacture process, geometric complexity has no influence on building efficiency, and therefore, no additional effort is required to design the molding structure, which makes 3D printing a very promising manufacturing technique [12, 22].

Despite these promising features, one of the main challenges for large-scale industrial use of 3D printing technology is how to eliminate or mitigate the structural distortion due to thermal effect during 3D printing processes, because it is a direct concern going against the main advantages of 3D printing: precision and as a single step manufacture. A typical 3D printing process is usually involved with melting materials into liquid or mushy-state and then rapidly prototyping it into various designed shapes. This process is a temperature drastic change process, and it causes material shrinkage, and various thermal-induced permanent deformations as well as residual stresses, which are induced when temperature changes during the printing process. This is not a simple thermal-mechanical process, and it is a complex interaction in and between the material layers in different material phases, and it governed by underlying process parameter variations and fluctuations [20].

In order to address this issue, various approaches or technologies have been used or proposed for distortion mitigation or compensation. The key for deformation compensation is to accurately predict deformation compensation allowance under a given set of 3D printing process parameters. It turns out that this is a very challenging problem that is largely open up to date. The most approaches to this problem are employing physics-based modeling to estimate the distortion al-
allowance based on simulations of the actual physical process of 3D printing. For example, the Finite Element Analysis (FEA) has been implemented to computer residual stresses and thermal distortions in various additive manufacturing settings, e.g. [13, 2, 17, 24]. By using coupled thermal, fluid flow and structure finite element models, researchers have studied the evolution of stresses and strains in various 3D printed products and their dependence on major process parameters including heat input and layer thickness, etc. [13]. However, these FEA modelings heavily depend on the accuracy and fidelity of the physical model adopted, and they are often limited by our lack of understanding of the actual physical process of a complex 3D printing process under investigation. Moreover, it is also very challenging to acquire the vast number of model parameters accurately in a actual manufacture process, and, on the other hand, the sophistication of the physical model also decreases its convenience to simulate the additive manufacturing process in order to compensate shape precisely.

In recent years, a number of data-driven methods in additive manufacturing have been developed based on existing machine learning strategies. For example, a novel data-driven approach that uses multi-scale multi-physics material models for additive manufacturing by self-consistent clustering analysis (SCA) has been developed in [23]. Li et al. [9] proposed an integrated framework that includes both classical numerical simulation and statistical method, such as, Gaussian process, and Bayesian calibration to predict thermal-mechanical field in additive manufacturing. These data-driven approaches use machine learning techniques to supplement the physics-based numerical modeling and simulation solutions. So far, to the best of authors’ knowledge, there has been little discussion, if there is any, on a fully data-driven approach in prediction of deformation or thermal compensation for additive manufacturing.

In this work, we have developed an entirely new data-driven approach to identify and quantify continuum/material distortion field in a 3D printed product in additive manufacturing. Our solution is solely based on the scan data on the spatial position of material points of the printed product and the designed configuration or designed data of the desired product, i.e. the initial or material configuration. The proposed continuum/material deformation finding algorithm or approach is closely interlinked to the coherent point drift (CPD) algorithm in artificial intelligence and statistics [14], and it may be viewed as a refined CPD algorithm for continuum deformation identification. The original CPD algorithm has demonstrated its dominance over most leading edge discrete point set registration methods. Nevertheless, to the best of authors’ knowledge, how to apply it to 3D printing technology has not been studied. On the other hand, the proposed material finding algorithm is a further development and an extension of original CPD algorithm to find continuum deformation or material registration in continuum media by aligning two stereolithography (STL) files.

There are five sections in this paper as follows: We start by recalling some key elements of original CPD algorithm in Section 2. Section 3 describes the proposed continuum/material deformation finding algorithm featuring the technique of aligning STL files. Numerical examples as well as the comparison of the numerical results with that of the CPD algorithm are provided in section 4. Finally, Section 5 provides the conclusion of this paper with summary.
2. Review of the Coherent Point Drift (CPD) algorithm

Since the resemblance between the continuum/material deformation finding algorithm proposed in this work and the original Coherent Point Drift (CPD) algorithm [15, 14, 19], for sake of self-containedness, we first briefly review the CPD method for non-rigid point registration such that the readers can become familiar with the idea, the principle, and the main techniques of the method and relevant literatures.

Originally, the Coherent Point Drift (CPD) method was first proposed and developed by Myronenko and Song in [14]. An example of registration problem which could be solved by CPD algorithm is illustrated in Fig. 2(a). The CPD algorithm first assumes that two point sets are given: (1) \( x_{m \times D} = (x_1, \ldots, x_m)^T \) is the first set of data for spatial point positions in deformation configuration, and (2) \( X_{N \times D} = (X_1, \ldots, X_N)^T \) is the second set of point positions in un-deformed configuration. After having those two point sets, the goal of this registration problem is twofold: (1) Align two point sets and find the correspondence pairs of points \((X_I, x_j)\), and (2) Find the non-uniform deformation map \( \chi \).

CPD algorithm cast the point set registration into a Maximum Likelihood Estimation (MLE) problem. For example, one can choose \( X_{N \times D} \) as the data points and \( x_{m \times D} \) as the GMM centroid. By considering a statistical MLE or optimization, in which one maximizes the probability for each data point by moving the GMM centroid, the spatial points may move to align (overlap) with the second set of points, i.e data points, in the undeformed configuration. That is to say, at the optimum, we match the points in two sets and obtain the correspondence pairs of points \((X_I, x_i)\). Finally, as in continuum mechanics, the non-uniform deformation map \( \chi : \mathbb{R}^D \rightarrow \mathbb{R}^D \) could be defined by the corresponding pairs \((X_I, x_i)\) with parameter \( \theta \),

\[
x_j = \chi(X_I, \theta)
\]  

(1)

The marginal probability density function for \( X_I \) used in the CPD algorithm is defined as follows,

\[
p(X_I) = w \frac{1}{N} + (1 - w) \sum_{j=1}^{m} p(x_j) p(X_I | x_j),
\]  

(2)

where \( j = 1, 2, \ldots, m \) (\( j = m + 1 \) is an added degree of freedom) is the index for the GMM centroid, and the index \( I = 1, 2, \ldots, N \) is for the data points. \( p(X_I | x_j) \) are conditional probability of data points with respect to each Gaussian distribution, while \( p(x_j) \) are membership probability for \( x_j \). Here we assume that each component probability shares the same value,

\[
P(x_j) = \frac{1}{m}.
\]  

(3)

In order to taking into account for noises and outliers, a uniform probability mass function \( P(x_{m+1}) = \frac{1}{N} \) is adopted for the mixture model. Then the final density function is a mixture model of original GMM and uniform distribution with parameter \( w \) which shows our prior belief of the level of noise in the data.
In the original CPD algorithm, the covariances matrix $\Sigma_j$ for each Gaussian distribution are assumed to be the same,

$$\Sigma_1 = \Sigma_2 = \ldots = \Sigma_j = \sigma^2 \cdot I$$  \hspace{1cm} (4)

where $I$ is the identical matrix and all covariances matrix $\Sigma_j$ are controlled by the variance $\sigma^2$. Then the conditional probability density function of $I$-th data point $X_I$ on $j$-th component in GMM is,

$$p(X_I|\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{D/2}} \exp\left(-\frac{\|X_I - \mathbf{x}\|^2}{2\sigma^2}\right),$$  \hspace{1cm} (5)

where the space is $D$-dimensional, in the registration cases, usually $D$ equals 2 or 3.

![Figure 2: The comparison between non-rigid registration method and material deformation finding method: (a) Non-rigid registration problem in computer vision vs. (b) Material deformation finding problem in 3D printing and continuum mechanics.](image)

In non-uniform deformation, the new position $x'_j$ of a GMM centroid $x_j$ is defined by the inverse of mapping function $\chi^{-1}$ with parameter $\theta$,

$$x'_j = \chi^{-1}(x_j, \theta)$$  \hspace{1cm} (6)

Thus, in essence, the CPD method is to use the maximum likelihood estimate to update the spatial position of GMM centroid $\mathbf{x} = \{x_j\}_{j=1}^m$ and covariance $\sigma^2$ such that it can converge to the data point set $X$. For instance, we may define the likelihood function as,

$$\mathcal{L}(X; \mathbf{x}, \theta, \sigma^2) = \prod_{I=1}^N p(X_I; \mathbf{x}, \theta, \sigma^2)$$  \hspace{1cm} (7)

that is

$$(\hat{\theta}, \hat{\sigma}^2) = \arg\max_{\theta, \sigma^2} \log\left(\mathcal{L}(X; \mathbf{x}, \theta, \sigma^2)\right)$$  \hspace{1cm} (8)

where $\hat{\theta}, \hat{\sigma}^2$ are the optimal values which determine both the GMM distribution and deformation mapping $\chi$. 

5
In actual computations, in order to find the point sets correspondence, CPD formulations use a so-called expectation-maximization (EM) algorithm, \[^5, 3\], which is an iterative or alternating method to find the maximum likelihood estimators on various incomplete data in statistics. More precisely speaking, the EM algorithm can use an iterative solution to find the maximum likelihood estimate or maximum a posteriori (MAP) estimate of interested parameters, where the statistics model heavily depends on missing data or the so-called latent variables. In EM algorithm, the first iterative step is a soft imputation on the missing data with a given probability distribution for all its possible values (E-step), and then the second step is to update parameters with the given imputed data (M-step) by maximizing or minimizing objective function \(Q\) that is found in the E-step. If we focus on the system parameters \(\sigma^2\) and \(\theta\), the objective function \(Q\) is defined as,

\[
L_Q(\theta, \sigma^2) = \frac{1}{2\sigma^2} \sum_{j=1}^{m} \sum_{l=1}^{N} p_{\text{old}}(X_l|x_j) \|(X_l - \chi^{-1}(x_j, \theta))\|^2 + \frac{N_pD}{2} \log \sigma^2,
\]

where \(x'_i = \chi^{-1}(x_j, \theta)\) is the transformed new GMM centroid locations;

\[
N_p = \sum_{j=1}^{m} \sum_{l=1}^{N} p_{\text{old}}(X_l|x_j)
\]

and

\[
p_{\text{old}}(X_l|x_j) = \frac{\exp \left(-\frac{1}{2} \left\| \frac{X_l - \chi^{-1}(x_j, \theta)}{\sigma_{\text{old}}} \right\|^2 \right)}{\sum_{j=1}^{M} \exp \left(-\frac{1}{2} \left\| \frac{X_l - \chi^{-1}(x_j, \theta_{\text{old}})}{\sigma_{\text{old}}} \right\|^2 \right) + (2\pi \sigma^2)^{D/2} \sum_{l=1}^{m} \sum_{i=1}^{N} w_i w_N \sigma_{\text{old}}^2}.
\]

Another way to think about EM algorithm is, after obtaining the system parameters \(\hat{\theta}, \hat{\sigma}^2\) in E-step, we may either move \(X\) to the deformed configuration, or move \(x\) to the undeformed configuration. By assuming that \(X^{n+1} = X^n + V^n \Delta t\) or \(x^{n+1} = x^n + v^n \Delta t\), where \(\Delta t\) is just a scalar parameter. The above objective function is also a function of the velocity field, i.e. \(L(X; x, \theta, \sigma^2, V)\). Thus, we may perform a maximization (M-step) of the objective function, i.e.

\[
\hat{V} = \arg \max_{\theta, \sigma^2} \log \left(L(X; x, \theta, \sigma^2, V) \right)
\]

Once we obtain the velocity field \(V\), we find the new position of GMM centroid, i.e. \(X^{n+1} = X^n + V^n \Delta t\), where \(V^n = \hat{V}\). Once this is accomplished, we come back to estimate \(\theta\) and \(\sigma^2\) again, and subsequently \(V^{n+1}\) and \(X^{n+1}\), until the convergence of the two point sets, i.e. \(X \rightarrow x\).

The essence of CPD algorithm is to ensure that the topological structure of data sets are preserved during registration. In order to address some computational issues of coherent moving, we may need add one prior term \(R\) to the MLE formulation and replace it into a Maximum a posteriori (MAP) estimation. The prior term encoded our prior belief of moving coherent and smoothness of displacement field. To do so, one can modify Eq. (9) by adding a regularization term \(R\) to obtain a constrained optimization problem,

\[
L_{Q\hat{R}}(\theta, \sigma^2) = L_Q(\theta, \sigma^2) + R(\mathcal{I}(x, \theta)),
\]

where \(\mathcal{I}(x, \theta)\) is an indicator function.
Specifically, the regularization term used in nonrigid CPD algorithms is

\[ R (\mathcal{F} (x, \theta)) = \int_{\mathbb{R}^D} \frac{|\tilde{v}(s)|^2}{\tilde{G}(s)} ds \]  

(13)

where \( G_{m \times m} \) is the Gaussian kernel matrix with element \( g_{ij} = \exp \left( -\frac{1}{2} \left\| \frac{x_i - x_j}{\beta} \right\|^2 \right) \). \( \tilde{G} \) is its Fourier transformation. \( \tilde{v} \) is the Fourier transformation of the velocity field. The integration is taken in the frequency domain with variable \( s \). This regularization term can first filter the high frequency part of the displacement field and then add penalty on it. By penalty the high frequency energy, this regularization term force the displacement field to put more energy on the low frequency domain and finally smooth the displacement field.

Figure 3: Material deformation finding for a deformed two-dimensional unit square: (I) Deformation of scanned data from the initial configuration to deformed configuration; (II) Deformation of designed mesh; (III) Topology connection between undeformed design mesh and scanned data in deformed configuration, and (IV) The exact solution for the material deformation finding problem.

3. The material registration in continuum media

3.1. Comparison between CPD and material deformation finding algorithm

In this work, a material registration in continuum media or a material deformation finding algorithm is developed to match the designed mesh of a 3D printed model and scanned material data
or mesh of the final printed product. We also assume that two point sets are given, as shown in Fig. 2(b): (1) \( \mathbf{y}_{m \times D} = (y_1, \ldots, y_m)^T \) is the first set of data for spatial point positions in deformation configuration, and (2) \( \mathbf{X}_{N \times D} = (X_1, \ldots, X_N)^T \) is the second set of point positions in un-deformed configuration. It should be noted that the spatial point set is denoted by \( y \) instead of \( x \) in material deformation finding algorithm. The reason for this modification will be demonstrated in the following paragraph.

The main challenge in this method is that there is no one to one correspondence between the nodes in two sets of meshes. A partial correspondence on the surface of the product may be established or identified by some geometric object identification methods, however, this can only be done for a group of nodes. In other words, even though we know the true deformation mapping function \( \chi \), we still could not find the image of a material point \( X_I \) in the spatial point set \( y \),

\[
y_j \neq \chi(X_I), \quad \forall j = 1, 2, \ldots, m
\]  

Therefore, the material deformation finding algorithm is to figure out the deformation mapping function \( \chi \) between two point sets \( X \) and \( y \) without the help of correspondence.

Besides the main difference above, there are also another three important differences between point set registration and material deformation finding problem. Firstly, in computer vision problem, both feature points in two point sets are obtained by feature extraction form image or triangulation data which means that each feature sets may contain noise. Our material deformation finding problem, however, use the exact design mesh as one feature set. Therefore, only the scan data set may contain noise data and the design data set is noise-free.

Second, in most cases, the feature points in computer vision application are only located on the boundary of the given model. On the contrary, the material deformation finding problem, as shown in Fig. 2, has feature points not only on the boundary but also in the interior domain.

Third, the objectives of CPD algorithm are to obtain correspondence between two point sets and generate the deformation function \( \chi \). On the other hand, the objective of material deformation finding algorithm is only to find the deformation mapping \( \chi \). One example of this difference is shown in Fig. 4, which is a 1d example of a straight line. The real deformation applied to the original red line is uniform stretching and translation to the bottom. The transformed line is the blue line. However, the locations of the two middle points in the blue line are chosen arbitrarily which means that the two middle points in the blue line are not corresponding to the two middle points in the red line. The registration result of CPD algorithm is shown in Fig. 4(a) where all the points are forced to match. The optimal result of material deformation finding algorithm that we wish to obtain is shown in Fig. 4(b) where not all GMM centroid \( y \) converge to a data point in \( X \). But the deformation mapping function generated by the movements of GMM centroid \( y \) is the real \( \chi \).

A summary of the differences is shown in the Table.
3.2. Motivation

Based on the comparisons on different aspects of the two methods listed in Table. [1] we conclude that the CPD algorithm may not be suitable to solve the material deformation finding problem. Therefore, a new method need to be proposed. As discussed above, the original CPD algorithm assumes equal isotropic covariances and equal membership probabilities for all GMM component. This assumptions fit well for the classical computer vision problems, where two point sets are corresponding features extracted from either 2D images or 3D models. In most cases, there is no underlying mesh where those feature points are selected from. In our application to 3D printing thermal displacement compensation problem, however, those two data points $X_{N \times D}, y_{m \times D}$ are vertices of the triangle mesh defined in the STL files. In fact, the mesh requirement is also not necessary, and the data points $X_{N \times D}, y_{m \times D}$ can serve as a meshfree representations of the given 3D...
printed models in different geometric configurations, i.e. spatial configuration as well as referential configuration.

In meshfree particle method [10, 11], the geometric state of a continuum body is generally represented by a finite number of discrete particles and the motion of the continuum body can be tracked by the motion of these particles. In the continuum problem domain, each particle represents a given fraction of the domain, by attaching a small volume of the domain, and the mass of each particle is proportional to the volume or area that each particle represents.

| Table 2: Pseudo-code for material deformation finding algorithm. |
|---------------------------------------------------------------|
| • Initialization: \( V = 0, \sigma^2 = \frac{1}{DNm} \sum_{j=1}^{m} \sum_{l=1}^{N} \| y_j - X_l \|^2 \) |
| • Initialize model parameters: \( w, \beta, \lambda, \tilde{W}_m \) |
| • calculate \( G : g_{ij} = \exp \left( -\frac{1}{2\beta^2} \| y_i - y_j \|^2 \right) \) |
| • Repeat the following E-step and M-step until convergence is reached: |
|   - E-step: Compute \( P \) and \( \tilde{P} \): |
|     * \( P_{ji} = \frac{\tilde{w}_m^{1-D/2} \cdot \exp \left( -\frac{1}{2\sigma_m^2} \| X_l - y_j + G(I, \cdot) V \|^2 \right)}{\sum_{k=1}^{m} \tilde{w}_m^{1-D/2} \cdot \exp \left( -\frac{1}{2\sigma_m^2} \| X_l - y_k + G(K, \cdot) V \|^2 + \frac{w}{1-w} \left( 2\pi \sigma_m^2 \right)^{D/2} m N \right)} \) |
|     * \( \tilde{P}_{ji} = \frac{P_{ji}}{\tilde{w}_m} \) |
|   - M-step: |
|     * Solve \( V \) based on \( (G + \lambda \sigma^2 d(\tilde{P}1)^{-1}) V = d(\tilde{P}1)^{-1} \tilde{P} y - X \) |
|     * \( N_\tilde{P} = 1^T \tilde{P}1, T = X + GV \) |
|     * \( \sigma^2 = \frac{1}{N_{\tilde{P}D}} \left( \text{tr} (y^T d(\tilde{P}^T 1) y) - 2 \text{tr} ( (\tilde{P} y)^T T ) + \text{tr} (T^T d(\tilde{P}1) T) \right) \) |

3.3. Material deformation finding algorithm

Inspired by meshfree particle methods e.g. [10], we assume that the membership probabilities for all GMM components and their corresponding covariances are proportional to the area that each scan data point assigned based on the underlying mesh. We can then define the weight \( \tilde{w} \) for each membership probability in \( y \) and variance \( \sigma^2 \) as

\[
\tilde{W} = (\tilde{w}_1, \ldots, \tilde{w}_m)^T \quad \text{where} \quad \tilde{w}_i = c \cdot \text{area}(w_i) \tag{15}
\]
where $c$ is a constant. To compute the weight vector $\tilde{W}$ and determine the value of $c$, we add one constraint to it: $\text{mean}(\tilde{W}) = 1$, which is

$$\frac{1}{m} \sum_{j=1}^{m} \tilde{w}_j = 1 \quad (16)$$

Therefore, the $j$-th component in GMM $p(X_i|y_j)$ and membership probability $p(j)$ should be modified to

$$p(X_i|y_j) = \frac{1}{(2\pi \tilde{\sigma}^2_j)^{D/2}} \exp\left(-\frac{\|x_i-y_j\|^2}{2\tilde{\sigma}^2_j}\right) \cdot \frac{\tilde{w}_j}{m}, \quad p(j) = \frac{\tilde{w}_j}{m} \quad (17)$$

where $\tilde{\sigma}_j$ is defined as $\tilde{\sigma}^2_j = \tilde{w}_j \cdot \sigma^2_j$.

Then the probability density function can be modified as

$$p(X_i) = w \frac{1}{N} + (1-w) \sum_{j=1}^{m} \frac{\tilde{w}_j}{m} p(X_i|y_j) \quad (18)$$

After recasting Eq. (2) into Eq. (18), the material deformation finding problem could be represented as a MLE and solved by EM algorithm as in original CPD algorithm. However, there is still one question need to be answered: Since there is no corresponding pairs $(X_i, x_j)$ between two point sets, why we continue using MLE to "match" two sets of points and what is the result of the "matching"?

The underlying assumption of this specific membership probability is that feature points are mesh-free representation of the model and the shape of the model should be independent of how to discretize it. Therefore, ideally, the GMM density function generated by any arbitrary mesh should be identical to the GMM density function generated by the uniform mesh. And the registration between a uniform mesh and an arbitrary mesh is conceptually identical to a registration problem between two same meshes which can be solved by CPD algorithm. For example, the GMM density function of a 2D model generated by different setting is shown in Fig. [5]. The Fig. [5] illustrates that the GMM density functions generated by uniform mesh and non-uniform mesh with specific membership probability for each member follow the same trend in the unit domain. However, GMM density function of the non-uniform mesh with equal membership probability has two sharp modes where mesh are dense which make it a poor representation of the true underlying density function. A numerical experiment of this model will be further studied in section 5.1.

By modifying the EM algorithm, we can write the objective function $L_Q$ in Eq.(9) as

$$L_Q(\theta, \sigma^2) = \frac{1}{2\sigma^2} \sum_{n=1}^{N} \sum_{j=1}^{m} p^{\text{old}}(y_j|X_n) \left\| X_n - \mathcal{J}^{-1}(y_j, \theta) \right\|^2 \cdot \frac{N_p D}{2} \log \sigma^2 \quad (19)$$

where $p^{\text{old}}$ is a conditional probability of each designed mesh nodal point to each member of
Figure 5: GMM density function: (a) non-uniform mesh with equal membership probability. (b) non-uniform mesh with specific membership probability for each member. (c) uniform mesh with equal membership probability.

GMM density function by using model parameters from the late iteration step,

\[ p_{\text{old}}^\text{old}(y_j | X_n) = \frac{\exp\left(-\frac{1}{2} \left\| \frac{X_n - \chi^{-1}(y_j, \theta_{\text{old}})}{\tilde{\sigma}_{\text{old}}} \right\|^2\right)}{\sum_{j=1}^{m} \exp\left(-\frac{1}{2} \left\| \frac{X_n - \chi^{-1}(y_j, \theta_{\text{old}})}{\tilde{\sigma}_{\text{old}}} \right\|^2 + c\right)} \]  \hspace{1cm} (20)

where \( c = (2\pi \sigma^2)^{D/2} \frac{w}{\sum_{j=1}^{m}} \frac{m}{N} \). In order to coerce GMM centroid to move in a consistent manner, we use a regularization term that is similar in CPD algorithm as

\[ E(V) = \text{Tr}(V^T G V) \]  \hspace{1cm} (21)

where \( G_{m \times m} \) again is the Gaussian kernel matrix and \( V_{m \times D} = (v_1, \ldots, v_m)^T \) is the coefficients matrix that may be interpreted as a pseudo-velocity, which determines the displacements. With this
regularization term, the objective function in Eq. (9) can be rewritten as
\[
\mathcal{L}_Q(\theta, \sigma^2) = \frac{1}{2\sigma^2} \sum_{n=1}^{N} \sum_{j=1}^{m} p_{old}(y_j|X_n) \left\| X_n - \chi^{-1}(y_j, \theta) \right\|_2^2 + \frac{N_p D}{2} \log \sigma^2 + \frac{\alpha}{2} \text{Tr} \left( V^T GV \right)
\] (22)

The transformation of GMM centroid can be defined as a generalized radial basis function which can be derive form Eq. (13) by using the representation theorem in the reproducing kernel Hilbert space (RKHS) (see [18]),
\[
\chi(y, V) = y + GV
\] (23)
Thus, the objective function in Eq. (22) can be written in matrix form as,
\[
\mathcal{L}_Q(V, \sigma^2) = \frac{1}{2\sigma^2} \left\{ \text{Tr} \left( y^T d(\tilde{P}^T 1) y \right) - 2 \text{Tr} \left( X^T \tilde{P} y \right) - 2 \text{Tr} \left( V^T G \tilde{P} y \right) + \text{Tr} \left( X^T d(\tilde{P} 1) X \right) \right\} \\
+ 2 \text{Tr} \left( V^T G d(\tilde{P} 1) X \right) + \text{Tr} \left( V^T G d(\tilde{P} 1) G V \right) + \frac{N_p D}{2} \ln(\sigma^2) + \frac{\alpha}{2} \text{Tr} \left( V^T GV \right)
\] (24)
where \( P \) is a \( m \times N \) matrix with element \( p_{old}(y_j|X_n) \), \( d(V) \) is a diagonal matrix with elements on the diagonal equal the value of a vector \( V \), \( I \) is a column vector where all the elements equal one.
\( \mathcal{L}_Q(V, \sigma^2) \) is a quadratic function of coefficient matrix \( V \). With the aim of updating \( V \) that minimizes the objective function in Eq. (24), we can set the derivative of function \( \mathcal{L}_Q(V, \sigma^2) \) to zero, which is
\[
\frac{\partial \mathcal{L}_Q(V, \sigma^2)}{\partial V} = -G \tilde{P} y + G d(\tilde{P} 1) X + G d(\tilde{P} 1) G V + \sigma^2 \alpha V = 0.
\] (25)
Then, we can find the value of coefficients \( V \) by solving the equation below,
\[
\left[ d(\tilde{P} 1) G + \sigma^2 \alpha I \right] V = \tilde{P} X - d(\tilde{P} 1) X
\] (26)
Similarly, we can update \( \sigma^2 \) as,
\[
\sigma^2 = \frac{1}{N_p D} \left( \text{Tr} \left( y^T d(\tilde{P}^T 1) y \right) - 2 \text{Tr} \left( \left( \tilde{P} X \right)^T T \right) + \text{Tr} \left( T^T d(\tilde{P} 1) T \right) \right)
\] (27)
To implement the above computational formulation, a pseudo-code of the material deformation finding algorithm is given in Table 2.

Remark 3.1. In the proposed material deformation finding algorithm, there are three free parameters: \( w, \lambda, \alpha \). Parameter \( w \) implies our prior belief of the level noise in the data set; And both \( \lambda \) and \( \alpha \) determine the level of motion coherence that one wants to enforce. We can tune all three free parameters until finding the best matching result.
4. Experiments and results

In this Section, we shall present two examples to demonstrate the accuracy, robustness, and improvement of the material deformation finding algorithm over the original CPD algorithm. The first example is a synthetic 2D problem in which the designed product is a two-dimensional unit square with a uniform triangle mesh and the scanned material data is a non-uniform or locally refined mesh on the deformed unit square. The material registration results are compared with the original CPD algorithm under same values of algorithm parameters. The second experiment is a real application of 3D printing example (see Fig. 1). The proposed material deformation finding algorithm is used to find the residual or permanent deformation of the 3D printed product in Fig. 1 due to temperature effect in an additive manufacturing process. Based on the predicted residual deformation, one can find the thermal compensation or compensation allowance, such that a new specimen may be printed on a compensated mold, and the final result is compared with the first printed specimen.

![Figure 6: 2D point set registration results of the proposed material deformation finding algorithm and the original CPD algorithm with different nonrigid deformation. (a) three different nonrigid deformation before registration. (b) the registration results of CPD algorithm. (c) the registration results of material deformation finding algorithm.](image)

4.1. Registration between uniform mesh and locally refined non-uniform mesh

As mentioned before, the main challenge in the determination of the distortion displacement field for a 3D printed structure component is that we only know the spatial positions of material
points in the current configuration, but we do not know the corresponding material coordinates in
the referential configuration. In order to find the displacement field, we have to find the material
coordinates of all spatial material points by material registration. To demonstrate the effectiveness
of material registration process based on the proposed material deformation finding algorithm,
in the first example, we choose a 2D specimen of a unit square as the shape in the referential
configuration, which is first meshed with uniform triangle mesh whose nodal points are serving as
the reference points, or data points. The uniform mesh in the referential configuration is shown
in Fig.3 II(a), which we also call as the design mesh in the referential configuration. Then we
create a locally refined non-uniform mesh by Delaunay triangulation algorithm [16], which is still
in the referential configuration as shown in Fig.3 I(a). After that, we apply non-zero strain or non-
rigid deformation to the square specimen, and it will deform into a near rhombic shape as shown
Fig.3 I(b). The spatial positions of the nodal points of the locally refined non-uniform mesh are
assumed as the scanned material data (see Fig.3 I(b)), which is usually known by scanning the 3D
printed product. In this way, the nodal point positions in the referential configuration i.e. Fig.3 I
(a) are their material coordinates, which are the exact solution of the material registration of this
synthetic problem, and they are unknown for any practical 3D printing problem. In this example,
we are trying to use the proposed material deformation finding algorithm to recover the material
coordinates for the scanned material points whose spatial positions as shown in Fig.3 I(b), based
on a set of reference or designed nodal points shown in Fig.3 II(a). Finally, we can compare the
numerical solution with the exact solution, i.e. Fig.3 III(a).

In actual numerical experiments, three different arbitrary nonuniform deformation are gen-
erated by using polynomial displacement field with random coefficients. Three different sets of
random coefficients are chosen with different levels of non-uniform deformation, i.e. with non-
zero strain distributions. Then we apply these non-uniform motion or deformation to the unit
square plate (see Fig. 3 I). In numerical experiment discussed below, we set the system parame-
ters’ values for the material deformation finding algorithm as $\beta = 3$, $\alpha = 2$, and $w = 0$ for all three
cases.

Figure 6 shows non-uniform material deformation registration results for three different pre-
scribed deformation cases. The left column shows three different experiments the design(red)
and scanned(blue) models. The middle column shows the registration results of CPD algorithm.
The right column represents the registration results of the material deformation finding algorithm.
Then we can interpolate the displacement fields in both $x$ and $y$ directions. The displacement dis-
tributions for case 1 are shown in Fig. 7 by color contour plots. The first row represents the value
of displacement in $x$-axis. The second row shows the value of displacement in $y$-axis. The right
column plots the assigned displacement field in both direction. The displacement field calculated
by CPD algorithm is shown in the middle column. The displacement field calculated by mate-
rial deformation finding algorithm is shown in the right column. It can be seen from Fig. 7 that
non-uniform deformation for all cases can be well captured by the proposed material deformation
finding algorithm, especially for the boundary points that define the shape of the domain, whereas
the original CPD method fails to register two data sets.

Also, in order to quantitatively compare the results of CPD algorithm and material deformation
finding algorithm, the registration errors of fours corner points in all three cases are calculated
based on Euclidean distance as illustrated in Fig. 8. For each level of non-uniform deformation,
our new material deformation finding algorithm always yields a significant better result than the CPD algorithm.

Figure 7: Displacement field of unit square for case 1. The first row represents the value of displacement in x-axis. The second row shows the value of displacement in y-axis. The right column plots the assigned displacement field in both direction. The displacement field calculated by original CPD algorithm is shown in the middle column. The displacement field calculated by material deformation finding algorithm is shown in the right column.

4.2. Additive manufacturing material deformation finding algorithm

The proposed material deformation finding algorithm is applied to 3D printing thermal distortion problem which is solely based on the scanned data of material point position of the printed product and the designed data of the product. The main idea is that: (1) The material displacement field is first obtained by material deformation finding algorithm. (2) The second step is to redesign the shape of product and the related 3D printing file or mold in order to obtain the final un-distorted printed product.

A flowchart of the material distortion finding algorithm with data pipeline is shown in Fig. 9. The starting point of this algorithm is two STL files. One of these STL files is the design mesh illustrated in Fig. 10 and another one is the scan mesh shown in Fig. 11.

Following the flowchart of the data pipeline, the next step is to create two feature point sets for the material deformation finding algorithm. Intuitively, we can choose all the nodes in two STL meshes as the feature points. Nevertheless, as discussed in the previous subsection, we only know
spatial positions of material points in the current configuration. However, we do not know their corresponding material coordinates in the referential configuration. In this particular case, there are 1,990,267 nodes in scan mesh and 44,868 nodes in design mesh, which means that it will be time-consuming to align all the original nodes directly. Therefore, in order to mitigate the above problem, we downsample the nodes by using a box grid filter. The size of the box grid filter is chosen as 2 mm in each direction for both meshes, which means that we enforce the condition that there is only one point in a size $2mm^3$ cube. If there are more than one nodes in a size $2mm^3$ cube, we will randomly choose one of the points in that cube. The results of the downsampled data are shown in Fig. 12 and Fig. 13.

In the next step, we start to use the material deformation finding algorithm to register two downsampled feature point sets with chosen parameters. In the simulation example reported in this paper, we choose the following values of the parameters: $\beta = 3$, $\alpha = 2$, and $w = 0.1$. In this example, the scan feature points are treated as GMM centroid, and design feature points are regard as the data points. Therefore, ideally, the scan feature points will move coherently to match the design feature points, and the displacement field can be interpolated by the displacements of the GMM centroid points. The registration result is illustrated in Fig. 14. The result of this registration shows that most of the feature points move coherently to the correct direction. However, there are still a few feature points that are misled by the noise in the data and could not find the correct corresponding positions.

To resolve this issue, we need to clean the GMM centroid displacements outlier before using them to generate the displacement field for the test specimen model. An robust outlier detection algorithm [7] are used to get rid of outliers and smooth the displacement interpolation field. The idea of this robust outlier detector is: the fluctuation of the of a given vector field is measured by a normalized residual fluctuation in its $k$ nearest neighbors [21]. In this method, the normalized
residual is defined in the following equation,

\[ r^*_0 = \frac{\left| \frac{U_0}{\text{med}(d_i) + \epsilon_a} - \text{med} \left( \frac{U_i}{d_i + \epsilon_a} \right) \right|}{\text{med} \left( \frac{U_i}{d_i + \epsilon_a} - \text{med} \left( \frac{U_i}{d_i + \epsilon_a} \right) \right) + \epsilon_a} \] (28)

where \( U_0 \) is the given vector value measured at one data point; \( U_i \) is the given vector value of its neighbors; \( \epsilon_a \) is the tolerance, and \( d_i \) is the distance from the data point to its \( i \)-th neighbor. After calculating the normalized residual \( r^*_0 \), it is compared with a threshold (default value is 2). If the value of \( r^*_0 \) is greater than the threshold value, this data point is marked as an outlier, and it should not be used in further calculation.

After detecting the outliers, we can calculate the displacement field \( \mathbf{u} \) by cleaned GMM centroid displacements

\[ \mathbf{u} = \mathbf{Y}' - \mathbf{y} \] (29)

where \( \mathbf{y} \) is the spatial location of the scan data nodes, \( \mathbf{Y}' \) is the location of the nodes of the scan data after material deformation finding, which, as shown above, is a good approximation of \( \mathbf{Y} \approx \mathbf{Y}' \).
Then, we can use the displacement value at GMM centroid to interpolate the displacements of other nodes. Both simple linear interpolation and \(k\)-nearest neighbors (KNN) regression \([1]\) are used for the interpolation. KNN regression is a non-parametric method, which interpolates the displacement field in a new configuration by its \(k\) closest training data in the feature space. Let \(N_k(y)\) denote the \(k\) values of \(y_1, \ldots, y_k\) that are closest to a new location \(y\) where need to interpolate.
Then the estimator is just the mean over the corresponding displacement value $u_i$ values,

$$ u(y) = \frac{1}{k} \sum_{i: y \in N_k(y)} u_i. $$

(30)

The motivation for using this estimator is that nearby observations is representative for the samples around a particular point $y$ [6]. Although this is usually not true, it may be a good approximation, if $u$ is smooth and the neighborhood support is small.

After finding the thermal induced displacement field, many compensation principles can be used to redesign the model. The simplest but still efficient way is to add an inverse displacement field to the original design mesh. The new added inverse displacement field will counteract the thermal induced distortion and finally achieve an neutral shape as prospect.

Both redesigned results of linear interpolation and KNN regression are illustrate in Fig. 15 and Fig. 16. The results of both redesigned mesh indicate that KNN regression is significantly better than the simply linear regression. These differences can be explained in part by the robustness of KNN regression, which does not require the regression function passing through all the data points.

In summary, the new thermal deformation compensation method uses material deformation finding algorithm to align feature points which are nodes in triangular meshes. Then after removing the outlier in displacement field, KNN regression is used to interpolate the displacement field at other points. Finally, redesign the model to compensate the distortion. This new compensation method starts with scanned data and design data files and ends up with a new redesigned STL file which could be use to print. It defines a new design-cycle to improve the quality of final produced 3D printing models. Future improvements may include design a more intelligent and maybe iterative way to use the thermal induced deformation field to redesign the model.

5. Conclusions

In this paper, we introduce a material distortion finding algorithm for predicting thermal compensation allowance in additive manufacturing.
In this data-driven approach, we only need to know the spatial positions of material points, i.e. the scan position in the current configuration, but not the corresponding material coordinates in the referential configuration or the designed configuration. By fully using the information contained in STL files, a unique weight associated with each feature point are calculated based on the area of triangle elements in STL files, which will lead to the spatial material points converge back to its original or material coordinates based on the maximum likelihood optimization. Finally, one can find the displacements as well as deformation of all the material points contained in the scan data. Different numerical tests have been carried out, and comparison study of results are also conducted. The material distortion finding algorithm developed here is illustrated to offer robustness to a greater degree with accuracy through quantitative evaluations, and it paved the way for registration to continuum manifold for significant applications in 3D printing technology.

**Acknowledgements**

S. L. would like to thank Ford Motor Company for awarding a Ford-University grant to support this research, and C. W. would like to thank a graduate fellowship from Harbin Engineering University supporting his Ph.D. study.
References

[1] Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician 46, 175–185.

[2] Arrazola, P., Özél, T., Umbrello, D., Davies, M., Jawahir, I., 2013. Recent advances in modelling of metal machining processes. CIRP Annals 62, 695–718.

[3] Bishop, C.M., et al., 1995. Neural networks for pattern recognition. Oxford university press.

[4] Campbell, T., Williams, C., Ivanova, O., Garrett, B., 2011. Could 3d printing change the world. Technologies, Potential, and Implications of Additive Manufacturing, Atlantic Council, Washington, DC 3.

[5] Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society: Series B (Methodological) 39, 1–22.

[6] Dudani, S.A., 1976. The distance-weighted k-nearest-neighbor rule. IEEE Transactions on Systems, Man, and Cybernetics, 325–327.

[7] Duncan, J., Dabiri, D., Hove, J., Gharib, M., 2010. Universal outlier detection for particle image velocimetry (piv) and particle tracking velocimetry (ptv) data. Measurement Science and Technology 21, 057002.

[8] Herzog, D., Seyda, V., Wycisk, E., Emmelmann, C., 2016. Additive manufacturing of metals. Acta Materialia 117, 371–392.

[9] Li, J., Jin, R., Hang, Z.Y., 2018. Integration of physically-based and data-driven approaches for thermal field prediction in additive manufacturing. Materials & Design 139, 473–485.

[10] Li, S., Liu, W.K., 2002. Meshfree and particle methods and their applications. Applied Mechanics Reviews 55, 1–34.

[11] Liu, G.R., Liu, M.B., 2003. Smoothed particle hydrodynamics: a meshfree particle method. World scientific.

[12] Mueller, B., 2012. Additive manufacturing technologies–rapid prototyping to direct digital manufacturing. Assembly Automation.

[13] Mukherjee, T., Zhang, W., DebRoy, T., 2017. An improved prediction of residual stresses and distortion in additive manufacturing. Computational Materials Science 126, 360–372.

[14] Myronenko, A., Song, X., 2010. Point set registration: Coherent point drift. IEEE transactions on pattern analysis and machine intelligence 32, 2262–2275.

[15] Myronenko, A., Song, X., Carreira-Perpinn, M., 2007. Non-rigid point set registration: Coherent point drift, in: Advances in neural information processing systems, pp. 1009–1016.

[16] Persson, P.O., Strang, G., 2004. A simple mesh generator in matlab. SIAM review 46, 329–345.

[17] Schoinochoritis, B., Chantzis, D., Salounitis, K., 2017. Simulation of metallic powder bed additive manufacturing processes with the finite element method: A critical review. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 231, 96–117.

[18] Schölkopf, B., Herbrich, R., Smola, A.J., 2001. A generalized representer theorem, in: International conference on computational learning theory, Springer. pp. 416–426.

[19] Wang, P., Wang, P., Qu, Z., Gao, Y., Shen, Z., 2011. A refined coherent point drift (cpd) algorithm for point set registration. Science China Information Sciences 54(12), 2639–2646.

[20] Wang, W., Cheah, C., Fuh, J., Lu, L., 1996. Influence of process parameters on stereolithography part shrinkage. Materials & Design 17, 205–213.

[21] Westerweel, J., Scarano, F., 2005. Universal outlier detection for piv data. Experiments in fluids 39, 1096–1100.

[22] Wong, K.V., Hernandez, A., 2012. A review of additive manufacturing. International scholarly research notices 2012.

[23] Yan, W., Lin, S., Kafka, O.L., Lian, Y., Yu, C., Liu, Z., Yan, J., Wolff, S., Wu, H., Ndip-Agbor, E., et al., 2018. Data-driven multi-scale multi-physics models to derive process–structure–property relationships for additive manufacturing. Computational Mechanics 61, 521–541.

[24] Yin, J., Zhu, H., Ke, I., Lei, W., Dai, C., Zuo, D., 2012. Simulation of temperature distribution in single metallic powder layer for laser micro-sintering. Computational Materials Science 53, 333–339.