High-throughput feature extraction for measuring attributes of deforming open-cell foams

Steve Petruzza, Attila Gyulassy, Samuel Leventhal, Jackson Baglino, Michael Czabaj, Ashley Spear, Valerio Pascucci

Fig. 1: Tracking of ligaments and junctions in an aluminum open-cell foam during incremental compressive loading from 0.0 to 2.0 mm. Red boxes highlight the changes in connectivity that we detect when ligaments fracture. The yellow boxes highlight part of the material drifting outside the field of view of during CT scans. Colors are mapped to the ids of ligaments and junctions.

Abstract— Metallic open-cell foams are promising structural materials with applications in multifunctional systems such as biomedical implants, energy absorbers in impact, noise mitigation, and batteries. There is a high demand for means to understand and correlate the design space of material performance metrics to the material structure in terms of attributes such as density, ligament and node properties, void sizes, and alignments. Currently, X-ray Computed Tomography (CT) scans of these materials are segmented either manually or with skeletonization approaches that may not accurately model the variety of shapes present in nodes and ligaments, especially irregularities that arise from manufacturing, image artifacts, or deterioration due to compression. In this paper, we present a new workflow for analysis of open-cell foams that combines a new density measurement to identify nodal structures, and topological approaches to identify ligament structures between them. Additionally, we provide automated measurement of foam properties. We demonstrate stable extraction of features and time-tracking in an image sequence of a foam being compressed. Our approach allows researchers to study larger and more complex foams than could previously be segmented only manually, and enables the high-throughput analysis needed to predict future foam performance.

Index Terms—Topological analysis, foam, features extraction, feature tracking

1 INTRODUCTION

Open-cell metallic foams and lattices are comprised of a network of interconnected ligaments. It is postulated that attributes such as the length of ligaments, connectedness of junctions, or distribution of pore size contribute to the mechanical performance of a foam. However, the relationships among these attributes and specific performance are not well understood. As foams are gaining more interest due to the plethora of manufacturing techniques, and simulation is increasingly able to reproduce performance characteristics, a data-driven approach becomes feasible to understand these relationships. However, high-throughput analysis of metallic foams is needed to populate this high-dimensional data space to analyze.

The state-of-the-art for analysis of metallic foams remains performing segmentation of features by hand, a time-consuming, laborious process that does not scale to the quantity of data sets, data sizes, and morphological complexities that need to be analyzed. Standard segmentation techniques struggle both to address the variety of structures present in open-cell foams, and to provide robust segmentation as the material is compressed to measure its performance characteristics. Drawing correspondences between junctions and ligaments becomes especially challenging as ligaments deform, self-contact, or fracture during compressive loading. A new approach is needed to address these challenges.

In this paper, we present a high-throughput, end-to-end workflow that provides robust and reliable segmentation of foams into junctions and ligaments. We introduce a new measure, the geodesic density function, that, combined with topological approaches, is able to identify junctions and ligaments consistently across time series images. We employ time tracking to disambiguate and explain the provenance of structures in the later time steps, when the foam has been compressed. Our system involves and guides the user in the analysis of a foam through careful use of visualization tools. Finally, we show that large-scale foams with thousands of ligaments, which were previously deemed too complex and costly to segment “by hand”, can now be analyzed with only minutes of user interaction. Our workflow increases the throughput for the analysis of open-cell foams, and marks the first step toward building data-centric models of foam performance.

2 MOTIVATION

Open-cell metallic foams and lattices are a class of structural-material systems that comprise a network of interconnected ligaments, resulting in a hierarchical structure [25]. These low-density, light-weight, load-bearing structures have been used in various multifunctional applications [4, 5, 23]. For example, they are able to serve concurrently as electrodes for energy-storage devices [62], as hosts for newly gener-
After manufacturing, each foam sample is mounted in a specialized in-work, we design and deploy a workflow for analyzing grayscale images. The foam models can be synthetically generated [2, 47, 57] or created directly from X-ray CT image data of an as-manufactured, uncrushed volume taken at each time step during the mechanical test. Similar image stacks can be generated from numerical simulations. For example, tensile open-cell foams, manipulating parameters of the investment casting process allows manufacturers to control the pore size (expressed by manufacturers in units of “pores per inch”) and relative density (which is governed by the shape and size of the ligament cross sections). For additively manufactured lattices, the geometry (including lattice spacings and topology) is controlled via the CAD model used to create the part. Discovering the process-structure-property relationships for a high-dimensional design space such as this requires generation of large amounts of data and observations, merit the implementation and use of the high-throughput feature extraction tool described herein. In this work, we design and deploy a workflow for analyzing grayscale images of metallic foams that were generated from experimental and simulated data.

For experimentally derived data, the images are collected as follows. After manufacturing, each foam sample is mounted in a specialized in situ load frame designed to enable three-dimensional imaging during mechanical crushing. During the mechanical loading, X-ray Computed Tomography (CT) data are collected incrementally [37] as shown in Figure 2. Data collected from the experiment are then post-processed to create a stack of grayscale images representing slices of the foam volume taken at each time step during the mechanical test. Similar image stacks can be generated from numerical simulations. For example, finite-element models of foam volumes can be generated and virtually crushed, and stacks of two-dimensional image slices can be extracted at different time steps, analogously to those extracted from experiments. The foam models can be synthetically generated [2, 47, 57] or created directly from X-ray CT image data of an as-manufactured, uncrushed foam [36].

We postulate that the following features of the foam will impact, to various degrees, the mechanical performance, including the onset and progression of deformation and failure:
- **ligaments**: length; orientation; cross section area; perimeter; and shape.
- **junctons**: volume; surface area; aspect ratio.
- **cells**: size; aspect ratio.
- **connectivity**: of junctions and ligaments; and of pores.

However, it is unknown how these specific features interact to govern the mechanical response and whether the features can be rank-ordered by their importance on the performance of the foam. This knowledge will be crucial to enable design optimization of the foams to meet specific performance requirements. The first step in enabling a data-driven design paradigm is collecting and analyzing large amounts of data and relating these properties to measured material performance.

**Prior state-of-the-art.** The grayscale images acquired from CT scans over the course of a compression sequence, in the past, have been used to generate animations that show the progression of compression. These animations are visually inspected to classify the behaviour of each individual ligament at each step [37]. Ligaments have been classified into one of four categories: those that have demonstrated brittle fracture with little plastic deformation, those that have exhibited only plastic collapse (such as buckling), those that have experienced a significant amount of plastic deformation before eventually fracturing, and those that have remained intact or mostly intact. This manual inspection and classification process is very time consuming and certainly not scalable to large volumes or time series. For instance, even a small example (with 82 ligaments) took domain scientists several days to segment and track only a dozen of ligaments. Hence, there is a need for solutions that can accelerate the segmentation and tracking.

**Challenges in segmenting CT foams.** One could imagine that the basic problem in creating an embedded graph structure to represent an open-celled foam could be addressed with standard approaches in image processing, such as medial axis transforms, skeletonization, or even topological connectivity structures. However, the neat mental model of “junction” and “ligament” collapses when faced with the reality of time-series of images of a foam being compressed. First, defects in real materials mean there is a continuum of morphologies between something that can be called a ligament and something that can be called a junction, as illustrated in Figure 3. For instance, when a normal ligament becomes too short, the junctions merge to form larger junctions; there is no clear demarcation between a thick ligament and merged junctions. Also, entire faces of the open-cell foam can be filled with material: what is an appropriate skeletal representation of such a shape? A primary contribution of this work is a novel measure of localized density to identify regions that act as junctions. Finally, as the foam is compressed, ligaments break, and previously separated structures now appear to have a ligament between them, precluding techniques that ignore time history and provenance of structures. In this work, we present a set of techniques that reliably and repeatably define and compute a junction and ligament model, and we track those structures throughout the time sequence.

3 RELATED WORK

**Skeletonization.** For the purpose of analysis and visualization of volumes, a simplified representative structure of shape is important. The exemplary curve-skeleton of the larger object is often interpreted to be a line simplification of the full volume, tracing the object’s center [53]. Various approaches to achieve this simplified skeleton include topological thinning [14], distance field based methods [11, 15], and potential field based methods [19, 63]. Previous distance field approaches to obtain a skeletonization have been based either on penalizing a Di-jkstra Shortest Path toward the boundary of the object from a core central line [11, 43], outward flux of gradient vectors from the center of the material [15], or computing the center line through level sets [31]. The interpretability of what defines a representative skeleton has led to numerous approaches of skeletonization as well as attempts at a formal definition [12, 19, 20]. Often skeletonization algorithms approach the problem by starting outside and working their way in, i.e., computing the center line of the full object through erosion [33],
thinning [7, 8, 13, 40, 42], and dilation from the object boundary [34]. A challenge when using these approaches to segment digital images is that the threshold determining the object interface completely determines the connectivity of the skeleton. There is no ability to estimate what components have been omitted. Another limitation of these approaches is that they do not relate the components of a skeleton to the void space outside the manifold of interest, for instance in connecting which portions of a skeleton surround a void or hole.

Of growing interest is the use of geodesics for constructing higher dimensional space summaries as lower dimensional path connections between points of interest or minimum spanning trees that convey significant structures within the higher dimensional shape. Previous approaches for determining these curvilinear structures have relied on computing distances from origin points using the Fast Marching Algorithm while minimizing an energy function or relevant measure [46] or iteratively moving from origin points to sink points through occlusion points of interest [18]. One such implementation used in neurite tracing determines geodesics as those propagated from the origin point that minimize a tubularity measure defined to be a spatial path as well as a curvilinear ‘thickness’ given by intersecting closed balls [6]. Performing image segmentation for holes, ends, and centers through mathematical morphology of geodesics has been of growing interest due to its success and intuitive physical meaning, allowing easy extension of geodesic morphological transformation techniques from two-dimensional image data into three-dimensional volumes or four-dimensional time dependent spaces [38]. Although skeletonization can be performed in geodesic space [9], our primary interest in it is a means of measuring properties from inside a topological space, specifically, the manifold pertaining to the metallic material.

Recently, approaches have begun to compute the curved skeleton through topological summaries such as Reeb graphs and merge trees based on persistence [16, 54]. The most successful methods employ the Morse-Smale complex, either directly on the function of interest, or on the signed distance field from a material interface. This approach has been used to construct the filamentary structure of the universe [51], represent bonds between atoms [10], trace lithium diffusion pathways [30], and extract the core structure of a porous material [29]. An advantage of topology-based techniques is they account for the impact of noise through persistence simplification [22]. We use the 1-skeleton of the Morse-Smale complex, as it enables reasoning about the stability of the extraction and provides a means of relating junctions and ligaments to the grains and open faces they surround.

**Feature Tracking** One central goal of our work is to track the skeletonized ligaments over the compression time series. Overlap between features in subsequent time steps is central to many approaches, and is used to categorize events such as continuation, creation, dissipation, bifurcation, and amalgamation [48, 49]. Such approaches rely on spatial overlap, and they have been extended to topological tracking of contours [3], for varying thresholds with Reeb graphs [21], space-time isosurfaces [56], or building a space-time function to track features defined on level sets [58]. However, compressed foams are imaged relatively sparsely in time, such that feature overlap is not reliable; with these techniques, moving junctions would instead be classified as appearing and disappearing. Basic nearest points have been used to track filamentary structures [29]; however, point correspondences were not used to infer object correspondences.

By providing a robust and concise feature description, skeletonization offers a new mean of feature tracking and matching. Through a hierarchical view of the computed spanning tree object, comparison becomes possible using isomorphic or nearly similar skeleton trees [17, 52]. This approach has also been employed in animation through computation of the centerlines, manipulation, and reconstruction of the original object [24]. Comparing objects based on their shape skeletons has also been used as a similarity measure for feature tracking over time [41] and feature matching following deformation [35, 61].

### 4 Generalized Geodesic Density

Often, structures of interest in an image may be determined by morphology rather than derived from image intensity. In the case of foams, ligaments and junctions may be locally indistinguishable in the image intensity (since the metal appears with the same intensity pattern in the images), and it is rather the local arrangement of material that defines a structure. Instead, a practical approach is to compute a field, such as the distance field, that encodes morphology, and then identify intensity features in the derived field. In the case of uniform radius tubes coming together at a junction, the intensity values of the distance field itself are not sufficiently sensitive to identify morphological structures. Instead, we introduce a new measure of localized density to be used to identify junctions. We first present basic geodesic density, which estimates the amount of material connected to and surrounding a point, and then improve the sensitivity of this measure further by generalizing the speed function used in geodesic distance.

#### 4.1 Geodesic Density

Let $M$ be a Riemannian sub-manifold of $\mathbb{R}^3$. The geodesic distance between points $a, b \in M$ is defined as

$$d_M(a, b) = \inf_{C_{ab}} \int_a^b dl,$$

where $L(C) = \int_a^b dl$, i.e., the length $L$ of the shortest path among the set of connected paths $C_{ab}$ containing $a$ and $b$. The geodesic ball of radius $r$ around a point $a$ is

$$B_M(a, r) = \{ b \in M \mid d_M(a, b) \leq r \},$$

that is, the set of points connected to $a$ with paths of length $r$ or less.

In our study of metallic foams, we utilized geodesics to characterize properties of points in the material in the context of their neighborhood connected through the material, rather than Euclidean distance. The primary motivation for this is that as the foam is compressed, junctions and ligaments that initially were separated may become abutting. However, the geodesic distances between points in the material are less affected by the deformation.

We introduce the geodesic density function $\rho'$ defined on a point in a manifold $a \in M$ and a radius $r$,

$$\rho'(a) = V(B_M(a, r))/V(B_M(r)).$$

The geodesic density is simply the ratio of the volume $V$ of a geodesic ball centered at $a$ with radius $r$ to the volume of a Euclidean ball with the same radius in the embedding space $\mathbb{R}^3$. The intuition is that density, $\rho = m/V$, is an amount of mass per unit volume. In our case, the submanifold $M$ represents metallic material, and by using the geodesic distance rather than Euclidean, the density corresponds to a measure local to the points in ligaments and junctions. In other words, adjacent, but structurally separated, ligaments and junctions are not considered when determining the density of a point.

#### 4.2 Identifying ligaments/junctions with geodesic density

We model ligaments as cylindrical “tubes” and junctions as intersections of at least three cylindrical segments. Intuitively, the density $\rho'$ at junctions is higher than along ligaments, as more material surrounds points in junctions. Consider a Euclidean ball in $\mathbb{R}^3$ with radius $R$ transected by a tubular segment (subset of $M$) with radius $r$, as illustrated in Figure 4. The volume of the tube with spherical caps is determined by the two radii, $r$ and $R$, as is the volume of an overall tube. We list the equations relating various structures with respect to a ball of radius $R$.

We wish to use $\rho'$ to distinguish between a tube, intersecting tubes, and disks. Given a distance field $d$ from the boundary of $M$, in Figure 4, the center of each has a value $r$ in $d$ (or $h = r$ in the case of the disk). These points are indistinguishable in the distance field, however, they have different values in $\rho'$, as is shown in the plots in Figure 5. The larger the radius $R$ of the “search ball”, the greater the difference between the values of the point in a tube, intersection, and disk. Figure 6 illustrates that the larger the search radius, the easier it is to define an iso-value that is localized to a junction.
4.3 Generalized Geodesic Density

As the material collapses on itself during compression, large merged areas become difficult to separate, even with the geodesic density. By modulating the local speed, it is possible to amplify variation in the image intensity that could correspond to the hairline interface between merged regions, improving the sensitivity of the geodesic distance. The junctions in Figure 8 need this improved sensitivity to be separable. A generalized geodesic distance simply takes into account varying speeds along paths between $a$ and $b$, and instead is defined as the time to traverse the fastest path between them. Let $s : M \rightarrow \mathbb{R}$ be a function that defines the speed at any point in the domain. Then the generalized geodesic distance between points is defined as

$$d^*_M(a, b, s) = \inf \{ L^*(C, s) \},$$

where

$$L^*(C, s) = \int_a^b \frac{1}{s(t)} dt,$$

i.e., the time $L^*$ of the fastest path among the set of connected paths $C_{ab}$ containing $a$ and $b$ given the speed function $s$. The generalized geodesic ball of time-radius $t_r$ around a point $a$ then becomes

$$B^*_M(a, t_r, s) = \{ b \in M || d^*_M(a, b, s) \leq t_r \}.$$

We introduce the generalized geodesic density function $\rho^*$ defined on a point in a manifold $a \in M$ and a radius $r$ with speed function $s$ as,

$$\rho^* = \frac{V(B^*_M(a, r, s))}{V(B^*_M(a, r, 1))},$$

that is, the ratio of the volume reachable from $a$ in time $t_r$ given speed function $s$, to the volume reachable in Euclidean space with constant speed of 1.

**Picking a speed function** The generalized geodesic density $\rho^*$ is more expressive, capable of adapting to the underlying intensity values in the image. We use the speed function to slow down propagation to avoid crossing cracks that appear as faintly diminished intensities in the material, to avoid gathering material from the other side in the density estimate. Therefore, our speed function is defined as $s : M \rightarrow [0,1]$, where $s(x) = 0$ if $x$ is a voxel that is unambiguously in the background, and $s(x) = 1$ if $x$ is a voxel that is unambiguously interior to the material. We allow the user to set two thresholds on the image intensity, a low threshold for background, and a high threshold for interior of material. The speed function parameterizes (and clamps if necessary) all intensity values to the range $[0, 1]$ with linear interpolation. Section 5.2.1 discusses more how a user selects the speed function.

5 Overview of Entire Workflow

The end-to-end workflow consists of four main stages: 1) creating junctions, 2) creating proto-ligaments, 3) time tracking and resolving ligaments, and 4) metrology. The first two stages of the workflow involve hands-on exploration of the data to determine thresholds for the feature extraction. Figure 7 provides an overview of the workflow, with numbered steps referred to in the following text. The green blocks in this diagram indicate user intervention to select parameter settings. The first, and most labor-intensive, part of the overall workflow is creating junctions. Once junctions have been identified, a skeleton is created of the material using a distance field and the Morse-Smale complex. The skeleton is combined with the junctions to create proto-ligaments.
ligament, structures that might be ligaments, but that need further verification. Next, time-tracking of the junctions is performed, and the corresponding time-tracking of the ligaments is inferred. The ligament tracking information allows classification and discarding of proto-ligaments that have no precursor in the prior time step. Finally, several metrics of interest are computed for each ligament over the time series.

5.2 Workflow for Creating Junctions

Given the time series of images $I_0, \ldots, I_n$, our goal is to generate geodesic density fields $\rho^*(I_0), \ldots, \rho^*(I_n)$ that have the property that junctions are identifiable as connected regions above a threshold $I$. Formally, the junctions are the super-level set components given a threshold in the geodesic density field. Furthermore, ideally, for each $i, j \in 0, \ldots, n$ a 1-to-1 map can be created between connected components of $\rho^*(I_i)$ and $\rho^*(I_j)$. As demonstrated in section 4, given an appropriate search radius and speed function, junctions can be distinguished with a threshold from ligaments in an idealized model. However, given the noise and morphological variability in real-world data, the radius selection, speed function specification, and threshold selection become a trial-and-error approach that relies on visual validation.

5.2.1 Generating generalized geodesic density

To generate a generalized geodesic density field that can reliably extract junctions as connected components given a threshold, a user first examines the last and then the first image in a time series. We have observed that in practice, the last time step represents the most complex scenario, where most of the ligaments are collapsed, and the material is compressed in a very limited portion of the initial volume. The last image gives an overview of what level of merging geometry happens, and how/if morphological structures are represented in the compressed state. The last image is used to identify the material and background values to be used to parameterize the speed function (from section 4.3), or if an additional processing step must be undertaken. The first image provides a pristine view of the ligaments, and is used both to select a search radius, and also visually verify that the chosen super-level set threshold creates a component at every junction.

Defining a speed function. During our thresholds selection, we want to understand which values in the volume under examination represent background and which instead represent material. The goal of using a generalized geodesic density function, as opposed to a regular geodesic density function, is to account for situations where a single threshold is not sufficient to determine which voxels are material and which are background. By controlling the speed, faint features separating morphological structures can be amplified, or in extreme cases where separate structures are not separated by image intensity, to fill holes by working on a distance field instead.

The input images are loaded in either ImageJ [44] or Paraview [1], depending on a user’s preference. A user navigates via slices to identify locations where the foam is pressed on itself. Certain CT scans have the sensitivity to represent hairline gaps between the structures, in which case the user picks a threshold (Figure 7, step 1) that represents “definite background” and another threshold that completely separates ambiguous regions (Figure 7, step 2) and can be interpreted as “definite material”. In other CT examples, and examples arising from voxelizing finite element simulations, there is often no intensity value that separates merged structures. In this case, we compute a distance field to the material interface threshold (see section 5.3), and select thresholds on the distance field, such that the morphological structures are separated, while retaining as much of the material as possible as interior. The thresholds selected for the last time step are reused for the entire image sequence.

Picking a search radius $R$. The relative ratios between the density value measured along tubes and at intersections increase with $R$, up to a limit of 2:3. Therefore, as $R$ increases, the selectivity of a threshold of $\rho^*$ increases as well; it is easier to pick a value of $\rho^*$ that includes a super-level-set component at each junction, while omitting unwanted points along ligaments. Two factors provide an upper limit to how large $R$ should be relative to $r$. First, the influence of a junction (or other thick material body) will extend at least $R$ from the ends of a ligament. With $R$...
too large, ligaments can erroneously count the material of the junctions at either end, making junctions harder to separate from ligaments. For the materials in this study, however, ligament lengths were much larger than the spacing between junctions, meaning that this concern did not impact the selection of $R$. The second limiting factor for the maximum size of $R$ is performance. For each voxel inside the material, $\frac{4\pi}{3} R^3 \rho$ potential voxels must be visited during the computation of the geodesic density function. Although evaluating $\rho^*$ is embarrassingly parallel, large $R$ values still prove prohibitively costly. In practice, we have found that a ratio of $\frac{R}{\rho} = 0.6$ consistently produced density fields from which it was simple to extract junctions using a threshold. For the materials studied, this corresponded to a search radius $R < 30$ voxels.

To select the search radius (Figure 7, step 3), a user loads the first time step, and similar to selecting the speed function thresholds, navigates through slices of the data. Using the visualization system’s measuring tools, the user estimates the radius $r$ of the largest tube to be considered a ligament. The value of $R$ is recommended to be $1.6r$.

**Selecting Junction Threshold.** The generalized geodesic density function $\rho^*$ is computed (Figure 7, step 4) given the user-selected parameters for the first and last time steps and written as raw files. A user loads these two files in Paraview as well as the two original images, and then uses the contour filter to investigate the density threshold to select in the context of a semitransparent material interface surface (Figure 7, step 5). Starting with the last time step, the lowest threshold value is picked that produces separated connected components for each junction. This same threshold is applied to the first time step where with the less cluttered configuration, a user is better able to visually verify that each junction has been extracted. Figure 9 illustrates this process. Checking that the number of connected components is the same for the first and last time steps is a strong indicator of stability of the junction extraction, however, not strictly necessary. For instance, in certain datasets from simulated foams, material is moved out of the field of view of the simulation domain, leading to fewer junctions in later time steps. Whereas in the vast majority of cases users reported that the original density function $\rho^*$ produced a satisfactory set of junctions, in rare cases the thresholds used to produce $\rho^*$ needed to be adjusted. For instance, in the case where no threshold exists that separates the morphological structures in the last time step, both the background threshold and the material threshold must be increased, and the density field regenerated. In the case that junctions are overly connected in the first time step, the radius $R$ needs to increase and a new $\rho^*$ must be computed. No case was encountered where $R$ was too large.

**Process Entire Time Series.** We reuse the four thresholds identified ($R$, background, material, and junction) and process the entire time series. If the background/material threshold selection required it, a distance field is computed for each image, and that field is used in place of the image itself for the rest of the workflow. The density field $\rho^*$ is generated for each image. The super-level set threshold identified during visual inspection is used to create an index volume from $\rho^*$, where each voxel is labeled either as background if it has value below the threshold, or with the identifier of its junction. A disjoint-set Union-Find data structure is used in a pass through the index volume to identify the connected components in the volume. Additionally, we generate boundary surfaces of each connected component using VTK [45] to enable rapid loading and visual verification in Paraview.

**5.3 Workflow for Creating Proto-Ligaments.** In a foam, the ligaments are portions of material connecting two junctions. For analysis of ligaments, the desired objective is to extract a lines in the medial axis of ligaments connecting junctions. Our strategy is to extract a skeleton representation of the foam and then identify the subset of its lines that form the ligaments. Our approach to skeletonization is based on computing the Morse-Smale Complex (MSC) of a distance field, simplifying it to remove spurious connections, and clipping the 1-skeleton using junctions. First, a threshold is selected to identify the interface between material and background. Next, a signed distance field is computed using this interface. The MSC is computed and simplified, and the ridge-like arcs entirely within the material are saved. The arcs are then clipped by the components of the junction, to create proto-ligaments. These are piecewise linear paths connecting two different junctions. Determining whether a proto-ligament corresponds to an actual ligament or to an artifact is left to the time tracking stage of the overall workflow.

**Estimating Material Interface Threshold.** The surface of the material is an isosurface of the image intensity. Picking the right intensity value to separate material from background is a first step to skeleton generation (Figure 7, step 6). The material scientists know, a priori, the relative density of the material for most samples. In this case, they supply a number that is the percentage of the foam volume that is material vs. the total enclosed volume of the sample, for the uncompressed (time = 0) foam. Foam images are available either as filling an axis-aligned rectangular domain, or as a cylinder inside the rectangular domain. In the former case, we use a threshold that simply includes the correct percentage of voxels. In the latter case, we found that manual intervention was needed to either set the threshold, or pick the subdomain on which the automated threshold could be computed. For both simulated and imaged data, the image acquisition parameters are kept constant so that the same threshold separates material from background in the entire time series. However, we found that for certain time series CT images, the average image intensity fluctuated between time steps. In this case, a new threshold was computed for each subsequent time step, by including the same number of high-intensity voxels as the prior step. In particular, given a threshold TR1 at time step T1 we compute the number of voxels in an iso volume at TR1 (i.e., number of voxels with value lower of equal to TR1) and find the corresponding threshold TR2 at time step T2 that will produce the same number of voxels in an iso volume at TR2. This approach uses the fact that no matter is created or lost during the time series.

**Distance field computation.** A distance transform is a field where where each voxel in a volume represents the shortest distance from that voxel to an object. In our case, the distance transform computed (Figure 7, step 7) for each voxel is the signed Euclidean distance to the surface of the material, with positive values inside the material and
Our feature tracking system does a forward trace of junctions and
5.4 Time Tracking Junctions, Resolving Ligaments
proto-ligament
the portion of the 1-skeleton that is interior to a junction (Figure 7,
newly created paths can be found between junctions due to the material
not necessarily indicate the existence of a ligament; in later time steps,
proto-ligament, because the existence of a path between junctions does
selves, are unstable, as shown in Figure 11b. Therefore, we remove
ridge-like arcs of the 1-skeleton cross between junctions. In the interior
of junctions, the exact location of the maximum, and the arcs them-
their pathways.
Proto-ligaments: connecting junctions with the 1-skeleton.
The ridge-like arcs of the 1-skeleton cross between junctions. In the interior
of junctions, the exact location of the maximum, and the arcs them-
several, are unstable, as shown in Figure 11b. Therefore, we remove
the portion of the 1-skeleton that is interior to a junction (Figure 7,
step 9). We define a proto-ligament as any remaining path connect-
ing two different junctions. We use the index volumes created during
the connected component sweep of the junctions workflow to query
whether a point on an arc is inside a junction. We use the terminology
proto-ligament, because the existence of a path between junctions does
not necessarily indicate the existence of a ligament; in later time steps,
newly created paths can be found between junctions due to the material
self-contacting. These are not to be considered in the final analysis.

5.4 Time Tracking Junctions, Resolving Ligaments

Our feature tracking system does a forward trace of junctions and
ligaments, using the resolved features of time step \( i - 1 \) to resolve
features at time step \( i \). It is a two stage strategy, first creating a mapping
between junctions, and then using that mapping to infer a mapping of
ligaments. Proto-ligaments in time step \( i \) are accepted as true ligaments
if a precursor is identified from time step \( i - 1 \).

Mapping Junctions. The relatively large time steps taken between
CT images restricts the kind of feature tracking that is possible. For
instance, overlap tracking fails because the junctions often move a
distance several times their own size. Instead, we employ a greedy
matching approach, where the largest, closest junctions are matched
first, followed by smaller, farther junctions (Figure 7, step 10).

Let \( J_i, J_{i-1} \) be the sets of junctions in time steps \( i \) and \( i - 1 \), respec-
tively. We wish to build a set of pairs \( \sigma = (γ_i, γ_i^\prime) \) uniquely mapping
between junctions of time steps \( i \) and \( i - 1 \). Each point \( p \) of each junc-
tion \( γ_i \in J_i \) is inserted into a spatial acceleration structure, in our case a
kd-tree \( K_i \). Similarly, each point of each junction \( γ_i^\prime \in J_{i-1} \) is inserted
into \( K_{i-1} \). Let ClosestJ(\( p \), \( K \)) return the identifier of the junction that
has the closest point to \( p \) in \( K \). We build a matrix \( m \) where for \( γ_i \in J_i \) and \( γ_i^\prime \in J_{i-1} \), the \( m_{i,j} \) element of \( M \) is

\[
m_{i,j} = \| p - γ_i \|_2 = \| \operatorname{ClosestJ}(p, K_{i-1}) - k \|_2 + \| q_i - γ_i^\prime \|_2 \]

Intuitively, the matrix \( M \) keeps track of how many times the closest
point \( q \) in the next time step to point \( p \) in the current time step occurred
for each possible pairing of junctions. Given \( M \), the pairing \( P \) is
computed by finding the highest valued element \( m_{i,j} \), pairing junctions
\( γ_i \) and \( γ_i^\prime \) into \( \sigma_{i,j} = (γ_i, γ_i^\prime) \), adding \( \sigma_{i,j} \) to \( P \), and then removing
row \( j \) and column \( k \) from \( M \). This greedy approach is repeated until
\( M \) is empty. We illustrate the results of the closest point queries from
every point in time \( i - 1 \) to time \( i \) in Figure 12a. In practice, \( M \) is very
sparse, is stored as a sparse graph, and the temporal matching is nearly
linear in the number of junctions.

Resolving, and Tracking Ligaments. Ligaments are resolved and
tracked inductively (Figure 7, step 11). We accept every proto-ligament
in the first time step \( i = 0 \) as a ligament. Let \( P \) be the set of pairings
produced by the junction tracking between time step \( i \) and \( i - 1 \). Each
proto-ligament \( ℓ \) has endpoints connecting two junctions in the same
time step. We accept proto-ligament \( ℓ \) in time \( i \) as a ligament, if the
junctions at its endpoints, \( γ_i, γ_i^\prime \), are mapped to the junctions time \( i - 1 \)
that are also connected by an accepted ligament. Intuitively, a proto-
ligament becomes a ligament if, mapping its junctions back in time, the
mapped junctions have a ligament in the previous time step. Therefore,
only connections initially present in the uncompressed foam can be
represented forwards by the time tracking: a desirable property when
the crushing can cause new connections to appear for material that is merely
adjacent. As a visual aide, and for debugging, we also provide users
with a visualization drawing a line segment between each vertex along
a ligament and the closest vertex on the tracked, matching ligament,
seen in Figure 12b.

5.5 Metrology
The quantities of interest to domain experts pertain mainly to the liga-
tments. The ligaments are available as a sequence of 3d points, and the
identifiers of the ligaments at each end point. Additionally the original
image, the distance field, and the density fields are sampled at each
point along a ligament. Understanding what metric impact material perfor-
ance relies on collecting a broad spectrum of information about the

Fig. 11: The 1-skeleton generated by the Morse-Smale complex places
a maximum (red spheres) in each junction (pink blobs) (a), with arcs
(yellow tubes) connecting them. This graph structure is used to derive
the connectivity of junctions. The locations of the 1-skeleton within disk-like
regions is not stable. In (b) the skeleton in consecutive time steps (red
for gray time step, blue for purple) shifts from one side of the disk to the
other, highlighting the need for stable junction extraction.

Fig. 12: We show the result of closest point matching (a) between junctions in consecutive time steps. For visual verification, we also produce closest point matching between tracked ligaments (b). The ligaments and their matches are colored according to feature id.
ligaments. The geometric realization of a ligament originates from the 1-skeleton of the MSC, which was computed using a discrete gradient approach. As a result, initial line segments on a ligament are aligned with axis orientation. We perform constrained iterative smoothing to remove these effects, and obtain a smooth line for each ligament. For each ligament, for each time step we compute:

**Length:** The accumulated euclidean distance between adjacent points on ligament

**Curvature:** For computing curvature $\mathcal{C}$ we consider a normal vector $\hat{n}_p$ to each point $p_i$ along ligament $\ell$ of length $h$. Adjacent points, $p_i$ to $p_{i+1}$, are considered to curve proportionally to the cosine similarity of these normal vectors. The curvature of the ligament is then the sum of all cosine similarities between adjacent normal vectors along the ligament.

$$\mathcal{C}(\ell) = \sum_{i=0}^{h} \cos(\mathbf{\hat{n}}_{p_i}, \mathbf{\hat{n}}_{p_{i+1}})$$

**Orientation:** Similarly, the orientation of a ligament defined to be the cosine similarity of the initial values of the 1-skeleton of a ligament to the $\hat{x}$, $\hat{y}$ and $\hat{z}$ axis vectors.

**Cross section area and perimeter:** We collect cross section areas and perimeters along several slices normal to the ligament along its path. As is done with curvature, to compute the cross section of a ligament we obtain the normal vector from a voxel on the 1-skeleton to it’s neighbor. With this normal vector and the perpendicular vector tangential to the ligaments 1-skeleton, we compute and take a cross sectional slice. Since multiple ligaments can be very near, potentially almost touching, it is necessary to identify which ligament was initially sliced. For this reason the ligament of interest is marked and once the slice is obtained we compute all connected components. The connected component then with the marked voxel is the cross sectional slice.

### 6 Evaluation

The workflow of the entire process (depicted in Figure 7) was evaluated using both simulated and imaged data. In particular, we considered a small simulated dataset of size $300 \times 300 \times 300$ voxels and four large experimental datasets with XY resolution $1024 \times 1024$ and different number of slices over the Z direction (see Table 1 for details). The evaluation was performed by four users who experimented with the process and stages of the workflow, and results were verified and validated at each stage.

**User Evaluation.** For a given foam data, the users performed all the interactive steps described in the previous section, and then visualized and evaluated the results. Of the four domain scientists involved in testing our system, two were considered experts (professors), as they collaborated during the development of the techniques, and two were undergraduates in materials science, previously involved only in sample generation and imaging. Users were guided only by a flow chart similar to the figures in Section 5 for running the tools, and a slide deck with screen-shots for guiding the user through Paraview’s interface. All software tools are linked with scripts, and produce state files for visual inspection in Paraview.

**User Feedback.** During testing of an early prototype, users needed several passes through the workflow to ensure the thresholds selected separated junctions of the compressed foam. After reordering the workflow to first look at the most compressed time step, and defaulting to use the generalized geodesic density, users reported that only occasionally would they have to complete a second pass of the entire workflow. For real CT image data, all users reported the ability to create satisfactory results with only a single pass for threshold generation. The greatest challenge was posed by simulated foams, where occasionally users required two or three iterations of the junction extraction, as the last time steps have heavily merged geometry. To ensure that the merged junctions could be separated required users to switch to computing the geodesic density on the distance field instead.

**Performance.** In Table 2, the computation times of the different analysis steps are reported for one simulated and one experimental dataset. The analysis were performed on a 12-core Intel(R) Xeon(R) Gold 6136 CPU @ 3.00GHz. The time for the users to select the various thresholds was under 5 minutes. In Figure 14 we report images of the features extracted in the four experimental foams, and we include a summary of the number of junctions and ligament extracted from each of them in Table 1. Users verified the results via visual inspection.

**Visual Validation.** Given that no ground-truth segmentation exists for the data being analyzed, all the metrics and features extracted were visually inspected and validated by the users at each step of the process. In particular, Paraview was used to visualize isocontours of the junctions in combination with tube visualization of the ligaments in the context of the material interface surface (as in Figure 14). Furthermore, coloring the ligaments by metrics (as in Figure 15) helped the user in validating the values of each. For example, in Figure 15, ligaments are either colored by inside/outside a junction (left), ligament length (middle), and orientation respect to the Z axis (right).

Finally, since ligament and junctions are tracked over time, we were able to observe the evolution of their metrics in order to detect changes in connectivity. In particular, a change in the connectivity of junctions and ligaments, or a drastic variation of cross section area over time can be indicators of onset of ligament failure (e.g., fracture). For example, in Figure 1 we highlight in the red boxes connectivity changes that have been reported by the tracking analysis. In particular, those changes correspond to ligaments that fracture at a certain time step during the
These results were obtained without supervision by nonexpert users. A synthetic dataset of size 300x300x300 and an experimental dataset molded to each other are unlikely to be separable, and will cause countered in the images processed so far, junctions that are completely total. A limitation of our approach is the assumption that junctions can of features at the boundaries represents a very small percentage of the features at the boundaries, also because for large datasets the number disappear from the boundaries. In our process, we currently ignore the out of the boundaries from one time step to the other (e.g., see yellow box, instead, is part of the material drifting in the yellow box, in the middle, the color of the tubes indicates the length of each ligament; and on the right, the color indicates the orientation respect to the Z axis of the ligament. compression. In the yellow box, instead, is part of the material drifting out of the field of view of the CT scan.

Finally, to demonstrate that our process can help classify the behavior of the ligaments, we observed the curvature value of each ligament tracked over time (see Figure 13). From these results, it is possible to quickly identify ligaments that are collapsing or fracturing during the compression (i.e., increasing curvature value over time, followed by disappearance). In Figure 13, we report four examples of ligaments with increasing curvature that are fracturing at different time/compression steps. In one particular case (i.e., second column from left), the ligament feature extracted (in red) is still contacting the two junctions even after the ligament separated (third image from top) because the ligament is still touching the junction and producing two connected maxima in the MSC. However, in the next time step, the connectivity is lost and the ligament is correctly classified as fully fractured.

### 6.1 Limitations

Both experimental and simulated data present certain challenges that our workflow is not currently addressing. For example, given the limited field of view of the CT scans, parts of the material can drift in and out of the boundaries from one time step to the other (e.g., see yellow box in Figure 1). This can cause junctions and ligaments to appear or disappear from the boundaries. In our process, we currently ignore the features at the boundaries, also because for large datasets the number of features at the boundaries represents a very small percentage of the total. A limitation of our approach is the assumption that junctions can be identified independently for each time step. Although not yet encountered in the images processed so far, junctions that are completely molded to each other are unlikely to be separable, and will cause a failure in their segmentation and tracking. Currently, our use of the generalized density does not leverage local material anisotropy, which could be added as a penalty term to the local speed function, for improved sensitivity to local morphology. Another limitation of this work is the reliance on the existence of maxima within each junction; one could imagine geometries, for instance a perfectly smooth T-junction, where no maximum is detected inside a junction defined by $\rho^*$. In this case, the persistence simplification might remove potential ligaments from the MSC 1-skeleton. Finally, computing the geodesic density function is expensive, practically limiting the search radius used, and discouraging experimentation with the speed function thresholds. We will look to accelerate this computation in future work. Finally, as foams become even larger and more complex, the visual validation parts of the workflow may become onerous and error-prone.

### 7 Conclusions

There is a high demand for means to understand and correlate the design space of material properties to the material performance metrics with respect to attributes of their features for open cell metallic foams. Prior to our work, the fine grain characterization of their features was accomplished via manual segmentation or with skeletonization approaches that may not accurately model the variety of shapes present in nodes and ligaments. In this work we introduce a high-throughput end-to-end workflow to segment foams into junctions and ligaments and track their behavior and properties over time. The task of tracking junctions and ligaments for each time-series of images is reduced to finding five thresholds: three to define the geodesic speed and radius, one for the material interface, and one for the junction threshold. Experiments performed by users on both simulated and imaged foams demonstrated the large benefits of this workflow, generating sets of thousands of ligaments and junctions with minimal user intervention. Furthermore, metrics collected from this process helped classify the behavior of ligaments over time, such as bending or breaking. The high throughput of our workflow is an essential step toward populating databases to be used to understand the parameters that regulate foam performance. In the future we plan to more fully automate the threshold selection (e.g., for cylindrical foams) and implement $\rho^*$ and other computational stages of the workflow using accelerators (e.g., GPUs). Furthermore, we will investigate utilizing neighborhoods in the junction/ligament graph to improve time tracking, and make it more robust to large-scale dislocations. Finally, we will investigate using existing junctions to postulate a segmentation for time steps where junctions are lost, indicating a catastrophic merging that is not separable using traditional approaches.

### Acknowledgments

This work was supported in part by NSF:CGV Award: 1314896, NSF:IP Award: 1602127, NSF:ACI Award:1649923, DOE/SciDAC DESC0007446, PSAAP CCMSC DE-NA0002375, NSF:OAC Award: 1842042 and NSF CMMI-1629660. Additional support comes from Intel Graphics and Visualization Institutes of XeLLENCE program.