Fast Detection of Material Deformation through Structural Dissimilarity

Daniela Ushizima*†, Talita Perciano*, Dilworth Parkinson‡
* *CRD, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
†*Berkeley Institute for Data Sciences, University of California, Berkeley, USA*
‡*Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
Email: dushizima@lbl.gov, tperciano@lbl.gov, dparkinson@lbl.gov

This work was partially supported by the Office of Energy Research, U.S. Department of Energy, under Contract Number DE-AC02-05CH11231.

Fast Detection of Material Deformation through Structural Dissimilarity

Daniela Ushizima*†, Talita Perciano*, Dilworth Parkinson‡
* *CRD, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
†*Berkeley Institute for Data Sciences, University of California, Berkeley, USA*
‡*Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
Email: dushizima@lbl.gov, tperciano@lbl.gov, dparkinson@lbl.gov

***Abstract*—Designing materials that are resistant to extreme temperatures and brittleness relies on assessing structural dynamics of samples. Algorithms are critically important to characterize material deformation under stress conditions. Here, we report on our design of coarse-grain parallel algorithms for image quality assessment based on structural information and on crack detection of gigabyte-scale experimental datasets. We show how key steps can be decomposed into distinct processing flows, one based on structural similarity (SSIM) quality measure, and another on spectral content. These algorithms act upon image blocks that fit into memory, and can execute independently. We discuss the scientific relevance of the problem, key developments, and decomposition of complementary tasks into separate executions. We show how to apply SSIM to detect material degradation, and illustrate how this metric can be allied to spectral analysis for structure probing, while using tiled multi-resolution pyramids stored in HDF5 chunked multi-dimensional arrays. Results show that the proposed experimental data representation supports an average compression rate of 10X, and data compression scales linearly with the data size. We also illustrate how to correlate SSIM to crack formation, and how to use our numerical schemes to enable fast detection of deformation from 3D datasets evolving in time.**

***Keywords*-pattern recognition; structure analysis; spectral methods; material inspection;**

I. Introduction

Scientific facilities that produce experimentally acquired data, such as light sources [2], will generate petabytes of data per hour by the year of 2020. Existing software tools for automated image analysis lack capabilities to process both current and future high-resolution image-based data coming from experiments. For example, much of the material failure investigation relies on visual inspection, a time-consuming task that hinders productivity and accuracy, particularly because the human vision quickly adapts to distorted signals among adjacent image slices.

Optical microscopy and 2D scanning have been the main imaging acquisition techniques for modeling deformation and damage of composites for many years. More recently, X-ray computed tomography (micro-CT) has offered new opportunities to perform microstructural analysis, particularly in investigations of fiber reinforced polymer (FRP). In fact, micro-CT (Fig. 1) holds the promise to improve the overall analysis of how complex FRP-based objects deform, damage, and ultimately fail [14].

This paper introduces automatic algorithms to evaluate FRP material deformation (Fig. 1) by calculating the relative importance of regions with respect to signal variation across different time steps. We frame key algorithms into distinct processes that can execute simultaneously, offering potential for significant speedup. Our pattern recognition pipeline for material deformation detection is designed to explore coarse grain parallelism, and consists of two sets of codes: a probe for dissimilarity among slices and algorithms for crack detection. Our proposed extension of the structural similarity index measure (SSIM) supports detecting degradation of material structure [13], and acts on image blocks that fit in memory and are independent. Concurrently, our crack detection algorithms seek micro-fractures and recover the volume of material failure. Together, these algorithms quickly lead the experimentalist to regions of the sample that contain distortions likely to be associated to material damage and/or image acquisition issues.

Our contribution is three-fold: (a) we design a scheme to compare local patterns of voxel intensities to deliver total variation curves of SSIM; (b) we deploy methods to detect micro-crack volumes from samples whose fiber bundles have a favored direction; (c) we lay out a processing construct that can explore parallelism to perform required computations.

The remainder of this paper is organized as follows. Sec. II describes the algorithms composing the proposed image analysis workflow. Sec. III presents the results obtained after applying the proposed algorithms to a 90GB of raw experimental data along with the computational performance analysis. Sec. IV and Sec. V provide discussions, conclusions and future works.

II. Materials and methods

Before real applications of new materials, such as ceramics, samples go through damage characterization to inform engineers about the structural dynamics. Time-lapse micro-CT imagery allows tracking 3D changes in material structure over time. The diagram in Fig. 2 indicates the data flow necessary for identifying damage processes from micro-CT images.

The proposed algorithms are part of an image analysis pipeline composed of two main science questions: (a)

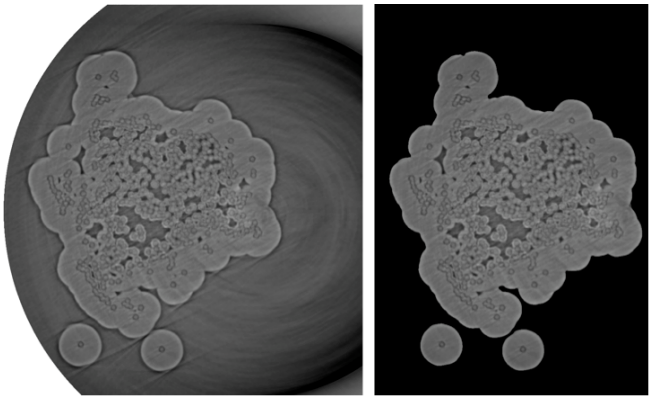


Figure 1. In situ X-ray μ-CT: reconstructed slice (left) and sample isolation (right) as in [11].

when does the damage start occurring? and (b) what is the volume of the crack opening? These are fundamental analysis requirements necessary to evaluate large datasets from structural dynamics experiments.

Tackling the question (a) requires three main tasks: 3D non-linear filtering, graph-based segmentation and spectral analysis. Specifically, we filter the data using our accelerated non-linear transformations in F3D [11], such as bilateral, median, and morphological operators in order to decrease noise and improve contrast. Next, we segment the composite into homogeneous regions using algorithms based on graphical models, such as Statistical Region Merging [4], and obtain a binary representation of the data. The last step takes advantage of the striped patterns of the data after vertical projection, followed by filtering and analysis in Fourier space, similarly to [12].

Our algorithm solution to the second question (b) also starts with raw data as input, but now using 2D slices of the microCT data. Each slice passes through a denoising step and then SSIM is calculated for the complete stack. The SSIM results provide inputs to design total variation curves in order to detect disparities throughout the image stack. The following sections describe details of the proposed algorithms used in this image analysis framework.

A. High-resolution 3D images

One way to reveal how new composites decay at microscale level is to deform samples by applying tensile loads using a displacement controlled loading system. This is the case of our datasets, which contain cross-sections of a ceramic composite under different compressive loads [1]. Using X-ray radiation to probe the material structure at micrometer scale, the LBNL ALS beamline 8.3.2 [2], [10] acquires samples at high-resolution. The reconstructed samples from the parallel-beam projection data are image stacks, as illustrated in Fig. 1.

We test our algorithms on these real experimental datasets consisting of silicon carbide-based composite samples, which are the result of materials load-bearing capacity tests under increasing strain. This paper refers to image stacks of 15GB [1], [11], with spatial resolution of 0.65μm, and under three different load (in newtons) conditions: 16.01N, 122.33N and 144.57N. Different phases of the material cannot be obtained through simple gray level cut-offs due to inherent artifacts, such as brightness variations across slices and noise, that may come from preceding image acquisition, reconstruction steps and/or sample composition.

The proposed algorithms consider 3D stacks of microtomograph images as input and output both the crack opening displacement volume as well as total variation curves during the evolution of the experiment for different loads. In order to access the sample at different scales as well as to compress the datasets, we design tools around hierarchical data formats as explained in the next section.

B. Terabyte-size image representation tools

One of the challenges of working with large image stacks is to maintain several copies of the data and at different resolutions. Despite the fact that microstructure analysis requires high-resolution data, much of sample processing starts with the identification of the sample bulk parts, for example, foreground and background, or mesoscale attributes, such as volume. Frequently, these multiscale representations are repeatedly computed or precomputed data versions are stored in the file system, lacking proper connection to the original data and/or connection to metadata associated to the experimental data collection.

Motivated by these demands, we calculate tiled multi-resolution pyramids at four different scales and store them in HDF5 chunked multi-dimensional arrays through BigDataViewer [6]. This plugin originally offers interactive arbitrary virtual reslicing of multi-terabyte recordings, so that the user can inspect the experimental data efficiently [5]. It can also be used to compress files, based on XML and HDF5, to allow encapsulation of terabyte-size image datasets, including metadata, and optimized access to multiple scales of the data, both for visualization as well as for processing. Other advantages of BigDataViewer formatting are: a) increased computing performance, b) decreased cluttering of the experimental archives, and c) potential for parallel I/O.

We build upon this work to deliver an image analysis pipeline that explores the multi-dimensional representation to provide a viable way to store and process large image stacks, and to enable the image facility users, such as those at the ALS [2], to access their data even when using their standard laptop. Sec. III demonstrates performance results regarding data compression.

C. Unstriped samples and crack opening displacement

Transverse CT slices parallel to the axis of the fibers produce striped patterns of the material. This image trans-

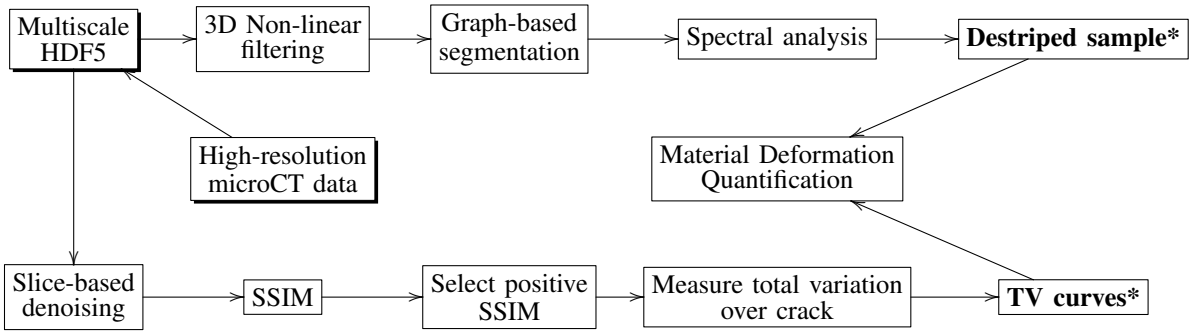


Figure 2. Image analysis workflow for high-resolution x-ray micro-tomography data: stack destriping and crack detection (* indicates proposed algorithms).

formation suggests an approach: by removing the fibers, whose intensity levels and local contrast descriptors are very similar to those at the microfractures, we can then recover the material crack opening displacement.

We propose destriping transformations based on spectral filters to focus on damage evolution on FRP. Relying on bandpass filters [12], our algorithm consists of suppressing structures that are parallel to the z-axis and smaller than 5μm, which corresponds to the fiber diameter. The striped patterns appear in the Fourier space as combined points of large amplitude, so those peaks can be detected and overcome. Moreover, the algorithm performs a shading correction by filtering in Fourier space.

By exploring the prior information, e.g., the fibers in our FRP samples display a particular direction, we can suppress stripes aligned vertically by compressing Fourier components along the fiber main direction. The unstriped image stack carries prominent information about cracks, which are easily binarized before the calculation of crack-opening measurements. Details about our segmentation algorithms for micro-CT can be found in [9], [10], [11].

D. Structural Similarity Index Measure

Image quality assessment aims at extracting quantitative measures to automatically translate perceived image interrelation into a structural similarity index measure or SSIM. Previous work [13], [7] estimates image degradations by exploring structural information from image intensity statistics between two images.

Different from previous work, we refactor SSIM to explore the structural information from two adjacent image slices, x and y. Next the SSIM algorithm separates the local variation of each slice into three components: luminance ($l(x,y)$), contrast ($c(x,y)$) and structure ($s(x,y)$). These components are calculated between consecutive slices of the stack in order to provide a similarity index $S(x,y) = f(l(x,y), c(x,y), s(x,y))$, which is symmetric, bounded and presents unique maximum. Using image intensity statistics to define each of the components, the resulting similarity measure, SSIM, is:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1)$$

where $C_1 = (K_1 L)^2$, for the dynamic range of the pixel values L and $K_1 \ll 1$ is a small constant, and $C_2 = C_1/2$. The statistics μ_x, μ_y, σ_x, σ_y and σ_{xy} are the local means, standard deviations and covariance for x and y, and are computed within a local 8 x 8 square window.

Building upon this work, we extend the SSIM concept to time-lapsed 3D image analysis and quantify image distortions across stacks to finally compare variations at different times steps of the experiment. This comparison entails the design of curves of total variation (TV) based on SSIM to find sample regions that contain distortions likely to be associated to material damage and/or image acquisition failure. Algorithm 1 describes the steps necessary to obtain TV curves, which include preprocessing of the cross-sections, followed by structural similarity calculation for two adjacent slices that produces a SSIM 2D matrix with same dimension as the cross-section. Finally, we calculate the integrated density, TV, normalized by the area of each z cross-section, for TV defined as in equation 2, where i and j are the 2D cross-section image indexes.

$$TV(z) = \sum_{i,j} SSIM_{i,j} \quad (2)$$

Fig. 5 illustrates TV curves, which enables to monitor image discrepancies throughout the 3D stack at different points in time, providing a quick feedback to the experimentalist with regards to sample deformation.

One of the main advantages of extending SSIM algorithm to stacks is the opportunity to parallelize the SSIM calculation at no communication of results between tasks since the

Algorithm 1 Total Variation curves from SSIM

Input: Original image stacks
Output: TV curves across z-axis for each time step

```
fiberDirection = vertical
for t = 0 to nsteps-1 do
    edge-preserving smoothing of Volume(t)
    for v = 0 to nslices-2 do
        x ← Volume(t,v)
        y ← Volume(t,v+1)
        s ← SSIM(x,y)
        if s_{i,j} <= 0 then
            s_{i,j} ← 1
        end if
        G_z ← integratedDensity(s) / area(s)
    end for
end for
```

slice computation is pairwise dependent. Considering a 3D stack at a specific time step, image blocks (sets of image slices) that fit into memory can be distributed across computing nodes where the SSIM is calculated independently. At a higher level, the computation of SSIM for stacks at different points in time can also be processed in parallel.

III. Results

The scheme proposed in Fig. 2 emphasizes the decomposition of computing tasks into two independent processes: process *A* using spectral analysis to detect the micro-crack volumes from unstriped samples, and process *B* using SSIM to compare local patterns of voxel intensities and extract total variation curves of SSIM.

Given the initial 90GB of raw experimental data, composed by 16-bit image stacks, we transformed the data into multi-resolution pyramids using BigDataViewer [6], [5]. The original data consisted of sets of *.tif* slices (2D cross-sections of 1,800 x 1,950 pixel), transformed into tiled multi-resolution pyramids at four different resolutions, 0.65, 1.3, 2.6 and 5.2 μm per voxel. The pyramidal representation implies no storage overhead for additional image tiles because BigDataViewer carries out the required indexing and transformation on the fly, based on automatically generated XML files. In fact, the total needed storage drops drastically with the BigDataViewer transformation, even though it allows retrieval of the images with the original high resolution and successively lower resolutions. Fig. 3 shows the scalability of this representation for increasing data sizes, providing an average compression rate of 10X.

Furthermore, the chunked multi-dimensional array scheme of BigDataViewer maximizes the performance of I/O operations at different resolutions of the image stack. The proposed algorithms for image processing at different image scales requires no communication between tasks, which

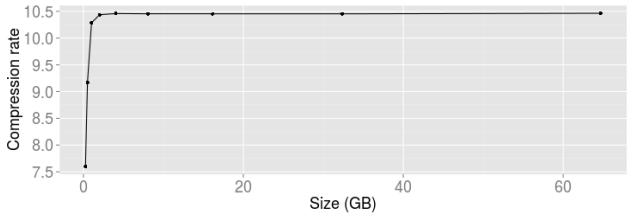


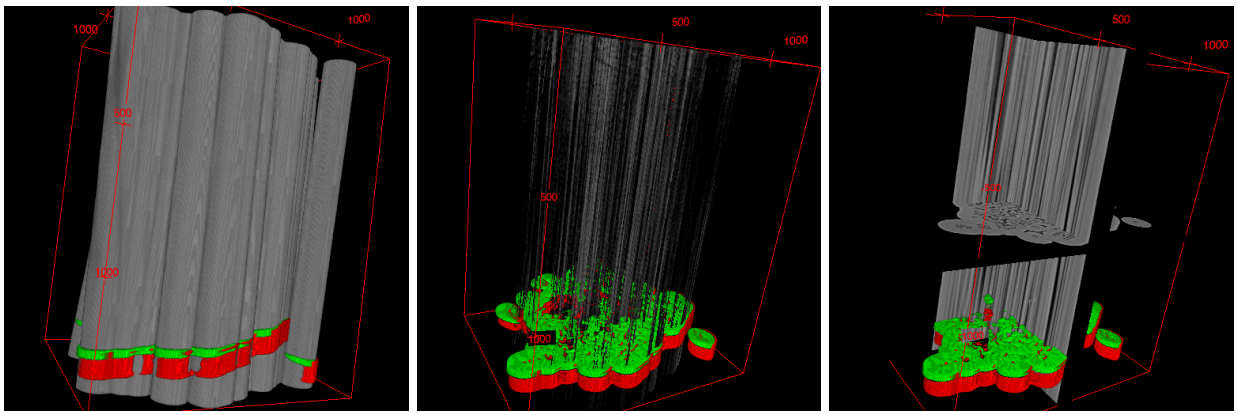
Figure 3. Scalability of the multi-dimensional representation using HDF5 with increasing data size.

makes the performance gain optimal for our image analysis pipeline.

Fig. 4 shows the results of executing the process *A* in order to recover the crack open displacement for three different steps of the experiment, and combining results in a single domain for visualization purposes. The volume rendering (gray level voxels) represent the sample before the micro-fracture since no crack was detected, the green surface indicates the detected crack displacement when the sample is under 122.33N, and the red surface is the complement to the evolving fracture, when the sample is subjected to 144.57N. Fig. 4(a) renders all time steps with no transparencies, while Fig. 4(b) emphasizes the crack opening displacements occurred in two different time steps and exposes the fibers removed by the destriping algorithm. Instead, Fig. 4(c) shows a cross-section of the sample under 16.01N and the detected crack opening displacements volumes for 122.33N and 144.57N. In addition, we illustrate the crack opening displacement volume, calculated for the intermediary time step (load = 122.33N) in Fig. 6.

The result of executing the process *B* appears in Fig. 5, in which we focus on a 1.5mm portion of the sample to illustrate how to probe for deformations. We calculate the average SSIM for each slice within the same stack, then we plot the resulting values in terms of the z-axis. We notice the existence of intervals of high dissimilarity, in other words, low average SSIM, along the samples with microfractures. This process takes usually less than a minute using a general purpose desktop, and it is much faster in comparison with other steps and can potentially prevent unnecessary explorations by process *A* when there is no crack. On the other hand, process *A* is essential to both precisely locate the crack and to calculate the crack open displacement since TV curves are restricted to probing the depth in which the microfractures initiated.

The proposed algorithms as well as all the employed image analysis tools were developed for Fiji [8], an open-source framework for image processing, mostly applied to biomedical images. All the figures were rendered using Fiji, except for Fig. 5, which uses R ggplot2 package.



(a) (b) (c)

Figure 4. Composite deformation under different loads: sample before crack formation in gray (16.01N), initial sample damage at 122.33N in green, and increased crack open displacement at 144.57N in red. (a) Rendering of three time steps with no transparencies, (b) cross-section of the sample under 16.01N, and (c) rendering exposing the fibers removed by the destriping algorithm.

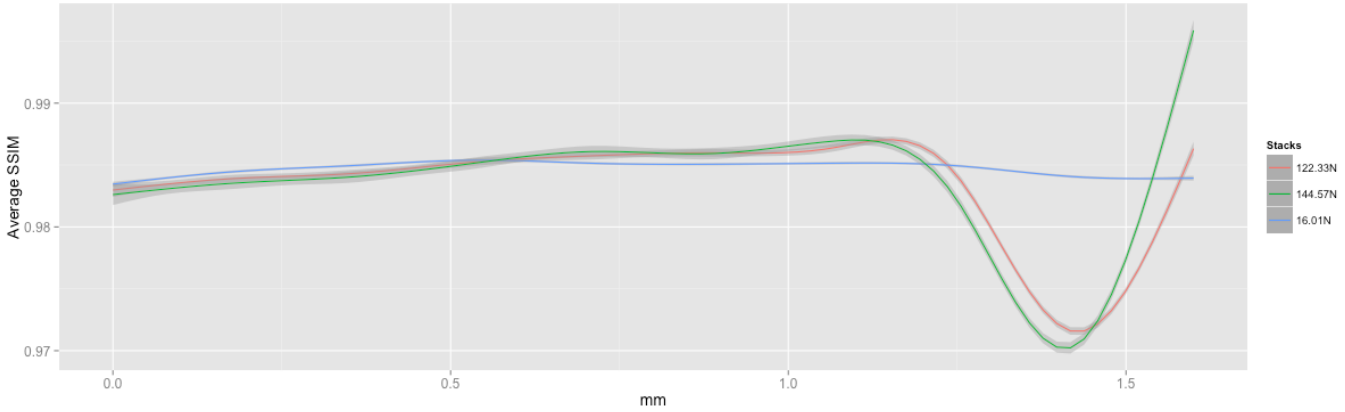


Figure 5. TV curves based on SSIM: valley indicates depth where the material cracks occur.

IV. Discussion

Microstructure characterization often requires evaluation of materials under extrinsic parameters, such as temperature and pressure. Before the utilization of composites, e.g., SiC-based ceramic FRP, into the design of turbines and aircrafts, several experiments are necessary to detect and evaluate microfractures.

Data representation is essential for fully inspecting high resolution images. Exploring the intelligent loading and caching scheme of BigDataViewer [5], we have enabled easy navigation of micro-CT data at multiple scales, avoiding aliasing artifacts at different views.

We introduced image quality metrics to probe image acquisition corruption and micro-damage, and illustrated the results using real datasets. Our preliminary tests showed that adjacent voxels tend to exhibit strong dependencies, which motivated the design of TV curves as spatially varying quality maps of the stack. These maps can provide information about the degradation of the image and/or sample deformations by pointing out the axial location of micro-fractures, as illustrated in Fig. 5.

This paper proposed software based on open-source tools to verify the existence of micro damage in 3D stacks of microtomograph images. It also proposed algorithms to detect the crack opening displacement as well as TV curves to quickly check and compare the evolution of the deformation when the material sample is under stress across time steps. We also introduced a pattern recognition pipeline that can explore parallelism to perform the required computations. The proposed use of the tiled multi-resolution pyramids stored in HDF5 multi-dimensional arrays also support analysis and processing of several resolutions independently and concurrently.

V. Conclusion and Future Work

We described an image analysis workflow for high-resolution 3D micro-tomography data composed of coarse-grain parallel algorithms. The proposed approach was tested

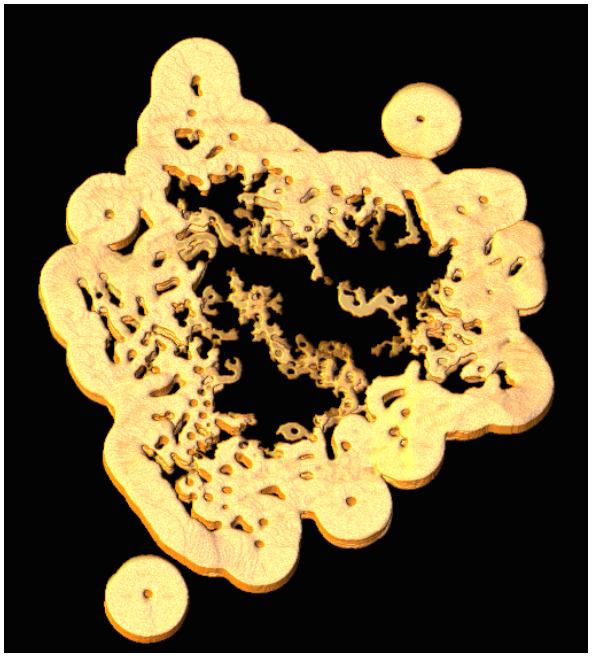


Figure 6. Rendering of automatically detected material crack opening.

on experimental micro-tomography datasets, allowing image quality assessment and material structural damage detection. We solved data representation issues by adopting a portable data format that provides efficient input and output, supports unlimited file sizes and has built-in and extensible compression facilities.

The presented analytical scheme offers several opportunities for parallelism, for example, by dispatching SSIM concurrently, which is calculated in independent image blocks. Also, TV curves and the crack detection calculations are independent, although these processes are complementary during the process of detecting material deformation from 3D temporal datasets. While this paper describes the design of the algorithms, the deployment of concurrent processing of several resolutions are to be delivered. Finally, each one of the processes individually provides a high scalability ratio with respect to the data size, given the low computing cost of SSIM and the potential for extending the spectral analysis using fast parallel approaches, such as ArrayFire [3].

Future directions include tests on larger datasets using extreme conditions of pressure and temperature, as well as statistical tools to show frequency of crack opening displacements, such as the thickness variation of the crack as exemplified in Fig. 6. New implementations might include comparisons with other crack detection methods, such as Hessian eigenvalues and percolation.

Acknowledgments

This work was supported by the Director, Office of Science, Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DEAC02-05CH11231 through the Early Career Research project and the Center for Applied Mathematics for Energy Related Applications (CAMERA). This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We also thank the LBNL Advanced Light Source (ALS) division for supporting the investigation and development of image analysis software tools to ALS user community. We are grateful to Dr. Ritchie and Dr. Bale for providing samples and domain information, and Dr. Pietro for specific information about SSIM.

References

[1] H. A. Bale, A. Haboub, A. A. Macdowell, J. R. Nasiatka, D. Y. Parkinson, B. N. Cox, D. B. Marshall, and R. O. Ritchie. Real-time quantitative imaging of failure events in materials under load at temperatures above 1,600C. *Nat Mater*, 12:40–46, 2012.

[2] J. Donatelli, M. Haranczyk, A. Hexemer, H. Krishnan, X. Li, L. Lin, F. Maia, S. Marchesini, D. Parkinson, T. Perciano, D. Shapiro, D. Ushizima, C. Yang, and J. Sethian. Camera: The center for advanced mathematics for energy research applications. *Synchrotron Radiation News*, 28(2):4–9, 2015.

[3] B. Kloppenborg. ArrayFire. *http://www.arrayfire.com*.

[4] R. Nock and F. Nielsen. Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1452–1458, 2004.

[5] T. Pietzsch. BigDataViewer. *http://fiji.sc/BigDataViewer*.

[6] T. Pietzsch, S. Saalfeld, S. Preibisch, and P. Tomancak. Bigdataviewer: visualization and processing for large image data sets. *Nature Methods*, 2015.

[7] G. Prieto, E. Guibelalde, M. Chevalier, and A. Turrero. Use of the cross-correlation component of the multiscale structural similarity metric (r* metric) for the evaluation of medical images. *Med Phys*, 38(8):4512–4517, 2011.

[8] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. Fiji: an open-source platform for biological-image analysis. *Nat Meth*, 9(7):676–682, July 2012.

[9] D. Ushizima, J. Ajo-Franklin, A. Macdowell, P. Nico, D. Parkinson, B. E.W, and S. J.A. Statistical segmentation and porosity quantication of 3d X-ray microtomography. In *SPIE Optics and Photonics*, volume 8135-1, pages 1–14, 2011.

[10] D. Ushizima, D. Morozov, G. H. Weber, A. G. Bianchi, J. A. Sethian, and E. W. Bethel. Augmented topological descriptors of pore networks for material science. *IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE Vis 2012)*, 18(12):2041–2050, 2012.

[11] D. Ushizima, T. Perciano, H. Krishnan, B. Loring, H. Bale, D. Parkinson, and J. Sethian. Structure recognition from high resolution images of ceramic composites. *IEEE International Conference on Big Data*, Oct. 2014.

[12] J. Walter. FFT filters. *http://rsb.info.nih.gov/ij/plugins/fft-filter.html*.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Processing*, 13:600–612, Apr 2004.

[14] W. Whitacre and M. Czabaj. Automated 3d digital reconstructions of fiber reinforced polymer composites. *American Institute of Aeronautics and Astronautics SciTech*, pages 1–18, Jan 2015.