PSE for Analysis of 3D Tomographic Images in Materials Science

Paulo Quaresma *, Pedro Medeiros *, Adriano Lopes † and Alexandre Velhinho ‡

*NOVA LINCS, Informatics Department, Faculty of Science and Technology, Universidade NOVA de Lisboa,
2829-516 Caparica, Portugal
Email: {pjq, pdm}@fct.unl.pt

† ISTAR-IUL, Information Science and Technology Department, Instituto Universitário de Lisboa (ISCTE-IUL),
1649-026 Lisboa, Portugal
Email: adriano.lopes@iscte-iul.pt

‡CENIMAT/i3N, Material Sciences Department, Faculty of Science and Technology, Universidade NOVA de Lisboa,
2829-516 Caparica, Portugal
Email: ajv@fct.unl.pt

Abstract—In the field of Materials Science, tomographic images play an important role in the analysis of composite materials. We present a computational environment that helps specialists in the field to carry out analysis and evaluation of samples of composite materials. This environment takes the form of a tailored Problem Solving Environment (PSE) and builds upon the SCiRun PSE. Its implementation is driven primarily by four major attributes: modularity, flexibility, interactivity and performance. Users can easily assemble networks of modules, with some of the modules being specifically designed for materials science analysis. These modules are flexible in terms of configuration, so yielding more flexibility to the setup of the networks, as well as in relation to the user interaction upon them once running. The implementation of data processing algorithms supporting critical modules rely on parallel programming. Furthermore, the quality of tomographic images under analysis is an issue of concern.

I. Introduction

RESEARCHERS in the field of materials science use X-ray micro/nanotomography (mCT) for studying composite materials, namely for 3D geometrical characterization of the material's constituent phases. The tomographic image that is reconstructed using specialized software corresponds to a 3D matrix, where each voxel in space is usually represented by an integer corresponding to its grey-level.

On that basis, it is important to provide materials science specialists with proper software to accomplish the research goals set. In particular, researchers are mainly interested on:

- Visualizing data in 3D;
- To perform different image processing operations in order to remove any artifacts present in the image;
- To exclude irrelevant objects to the ongoing analysis;
- To perform image processing operations that label each of the distinct objects under consideration;
- To obtain geometric information that establishes a statistical description of the entire population of objects under consideration.

In this article we describe a framework for building flexible environments for the analysis of tomographic images of composite materials. Besides the normal operations we may expect to use in tools of this category, this framework is mostly concerned with usability and data quality issues that materials science specialists might face. They are:

- Modularity and flexibility, as the system must support an easy way of specifying the processing steps, and should allow to easily perform testing and reconfiguration tasks;
- Interactivity, in the sense that individual processing and visualization operations should be carried out faster since specialists want to see the outcome of those operations as quickly as possible, and also to allow a smooth steering of the computations;
- Tomographic image quality, since it should not be taken for granted that all images will show high contrast.

The organization of the paper is as follows: Section II presents related work that has been developed in the area of computational environments for analysing scientific data. Then, in Section III, we introduce a framework alongside guidelines to build a computational environment to process and analyse scientific data, followed in Section IV by an implementation with focus on data collected from material science experiments. In order to validate our proposal, we discuss a case-study in Section V, in particular concerning tomographic images with low contrast, so difficult to process, and finally Section VI wraps up with conclusions.

II. Related Work

Computational environments to process tomographic images can broadly be split into two major categories: environments that allow users to apply processing algorithms to tomographic images on a one-o-one basis, that is, with the simple paradigm *read-transform-visualize* in sight, and the so-called visual programming environments, more friendly but complex, which allow users to set up a network of processing modules via graphical deployment in a canvas, with related computations following the data–flow model [1].

As inferred from above, visual programming environments allow specialists to specify a sequence of processing steps by choosing a set of modules available in a menu, including obviously reading tomographic images, and interconnects them. An example of those is the commercial software Avizo/Amira [2].

On the other hand, one can prefer to use the other category of environments, like the image processing and visualization software ImageJ/Fiji [3] or Paraview [4]. Worth pointing out that we can always follow the route of developing dedicated software, like the case of spam mentioned in [5].

In respect to visual toolkits mentioned above that rely on the data-flow model, sometimes referred to as PSEs due to its usability in various scientific areas, it is clear that a major advantage they present is that they can be tailored to the specific needs of a particular scientific area, yielding to dedicated environments. One example is SCIRun [6] from the University of Utah, USA. The toolkit has been very successful regarding the development of dedicated PSEs. For example, it is the case of BioPSE [7], which is specifically tailored for running bio-electric field simulations on top of SCIRun.

III. FRAMEWORK

Given the information provided by practitioners in the materials science field, our understanding is that we should have an integrated software solution, embracing both image processing algorithms and visualization capabilities, but underlying a clean and easy-to-use approach. On that basis, the proposed solution will address primarily the following requirements:

- Open-source desktop solution but providing users interactivity and computational steering;
- Processing of tomographic images, including the ones with low contrast;
- Availability of various processing algorithms, even the complex ones requiring higher computational resources;
- Providing adequate data formats in accordance to the processing operations of concern.

We borrow the idea from the concept of PSE, sustained by the data-flow model [1], upon which SciRun is a prominent example. Hence, our solution provides four distinct categories of modules: data readers, filters, mappers and renders. Fig. 1 depicts the processing model we advocate.

Specialists will have at their disposal such modules to create networks, that ultimately will solve their problems. The networks will be managed by the specialists themselves. This includes the setting of control parameters and of both data and images to/from modules via input/output ports.

IV. IMPLEMENTATION

All the developed modules were built on top on SciRun. Notice that the interactivity and steering requirement is delivered by SciRun. Next, we will introduce new implemented modules, yielding to an open-source solution that works on desktops.

A major concern is that tomographic images showing up low-contrast between matrix and particles are challenging to identify and characterize objects – let us focus hereafter on particles. Common approaches sometimes fail to to so and some take too much time to deliver results. That is why we have taken a careful approach while designing those modules that are related the most. For example, modules belonging to the category *Filter* have been implement using OpenMP or CUDA, so a parallelization approach targeting both CPUs and GPUs. Operations that do occur at voxel level will take advantage of data parallelization.

In respect to visualization functionalities, we take advantage of native SciRun visualization modules, mostly for general 3D visualization. But for specific purposes, like visualizing and analysing particle features, specialists are able (i) to automatically launch external viewers and, importantly, (ii) to use a new 2D visualization module to check features on a image plane basis, regardless of its orientation in the 3D space. Also, for a better understanding, specialists can playback the outcomes of image operations that were applied, in sequence.

In relation to image operations, and among the various modules that have been implement, there are some operations that deserve to be singled out. They are: edge detection, segmentation, erode and dilate, and crucially particle identification and subsequent characterization.

Edge detection. This operation basically creates conditions to correctly identify particles. Examples of filters that relate to this task are *Unsharp* (mask to unsharping to enhance high-frequencies like boundaries), *Gradient* (first derivatives), *Laplacian* (second derivatives to enhance tiny boundaries), *Sobel* (Sobel derivatives to give direction of intensity variations) and *ZeroCrossings* (location based on 2nd derivatives).

Segmentation. The goal in this operation is to highlight particles within the raw image. This is carried out via the *Thresholding* filter: Assuming that we have a raw image defined in a gray scale, by applying the filter we get a black-and-white image. Also, we can chose to apply bi-segmentation, meaning that we end up with black, or white, or unchanged voxels. Crucially, this operation requires setting critical cut-off levels, usually inferred with the help of image histograms.

Erode and Dilate. These two operations working together help to achieve particle separation. It happens when somehow the boundaries of particles touch to each other, that is, it seems there are contiguous voxels but belonging to different particles. Notice that the given image is already in black-and-white. *Dilate* implies enlarging the border of a particle (white to black), whereas *Erode* is the converse operation. When there is a sequence *Erode* then *Dilate*, it is called *Open* operation. The reverse sequence is called *Close* operation.

Particle Identification. This crucial task implies the use of various filters supporting a bi-segmentation process, which is even more critical when the raw tomographic images show low contrast between matrix and reinforcements. It also relies on two modules: *ParticleLabelling* and *PoissonReconstruction*. The first one is based on a labelling algorithm [8] but implemented with OpenMP or CUDA, (there are two versions available, meaning it is up to the specialist to decide which version is going to be used) yielding to a fine-tuned parallel implementation to reduce execution times. The second module uses
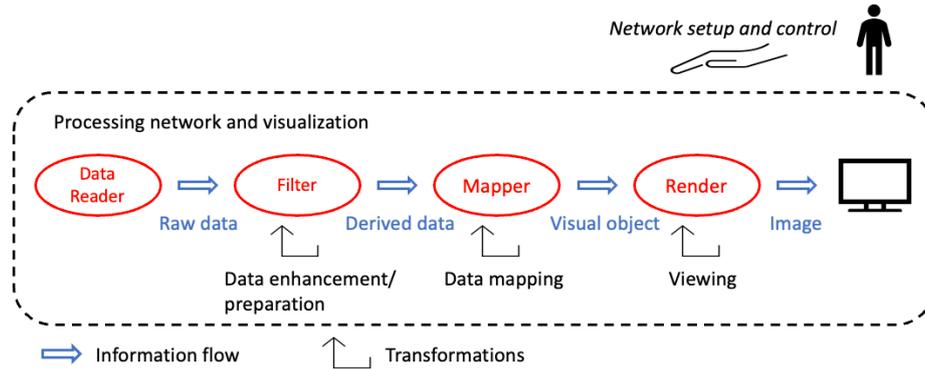
Fig. 1. Processing model based on the data-flow visualization paradigm.

the Possion surface reconstruction algorithm [9] available in the Point Cloud Library (PCL) library. (https://pointclouds.org) While the first module produces sets of interconnected voxels, (parts of potential particles) the second module takes a cloud of points as input (those voxels) and properly reconstructs the surfaces of the particles. Then, as output, polygonal meshes are generated to depict the reconstructed surfaces.

Particle Characterization. Once particles have been geo-metrically identified, specialists still have to further evaluate the outcome as it is presented to them. Notice that we may end up with clusters of particles or even *fake* particles. But in the end, the final decision about accepting or rejecting a particular particle rests on the specialists. The outcome will be a set of particles of interest and their characterization – the location within the sample, the geometric profile such as volume, area and bounding box, among other similar concepts.

Worth pointing out that the way data is organized affects the performance of applied algorithms. That is why it has been also implemented a set of rapid but robust data formats converters. Then, in a particular situation, specialists will decide which ones to use in order to achieve better efficiency and performance. Also, because the visual appearance of particles is helpful in the characterization process, specialists have at their disposal various visualization functionalities.

V. EVALUATION

In order to validate the extended PSE, we discuss now a case-study concerning samples of aluminum as the base material (matrix) and tungsten carbide as reinforcements but showing low contrast between them. That is, if we were to draw an histogram of densities, it will not show two clear peaks – one corresponding to the matrix and another one to the reinforcements – as we were expecting to obtain in a clear bi-segmentation process.

The samples were collected at the European Synchrotron Radiation Facility in Grenoble and were defined in a regular grid. For evaluation purposes, we use a subset corresponding to a uniform 3D cube lattice of dimension $[512 \times 512 \times 432]$, with each voxel corresponding to one μm^3 approximately [10].

Looking at the raw tomographic images, they show low contrast between the base material and reinforcements, and

contain various porous. (See Fig. 2) Furthermore, information gathered during the collecting process hinted that the particles were showing a cone-shaped, convex geometry, they could be broken, and the average size is about 35 μm.
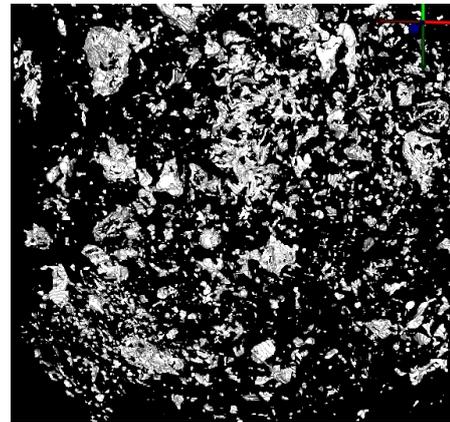

Fig. 2. Glimpse of a raw tomographic image prior to any processing.

Overall, the goal is to identify particles inside the sample an then to characterize the ones of interest. The sequence of operations works as follows:

1) Removal of porous;
2) Increasing contrast between particles and base material;
3) Particles labelling;
4) Particles detection;
5) Particles characterization.

The initial two operations are mostly supportive of the parti-cles labelling process. Hence, and given the modules available, a typical workflow to accomplish the tasks mentioned can be split into three sequential stages: particles labelling, particles detection and finally particles characterization. In the follow-ing we will provide further details about these three stages.

A. Particles labelling

The goal here is to figure out potential locations of particles in the volumetric sample. It starts by smoothing the raw data, that is, reducing the noise in the tomographic image and then,

in sequence, applying band pass filters, labelling the image, followed again with enhancement using pass filters.

For example, Fig. 3 shows a circular air porous that is going to be removed by first painting its interior with the colour of the matrix so once bi-segmentation is applied later on, it will be converted into matrix. At this point we are able to get an initial identification of reinforcements.
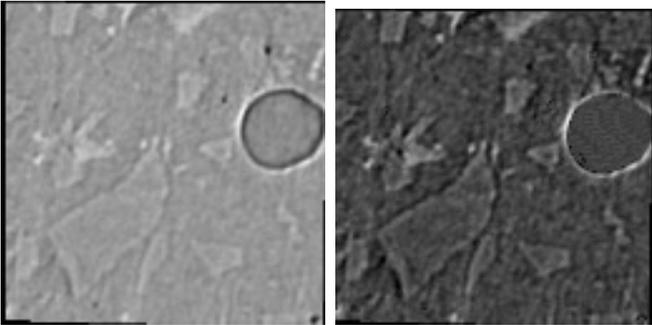

Fig. 3. Circular air porous (left) that will be removed once bi-segmentation is applied, but only after pre-painting its interior as matrix (right).

Then, it follows a bi-segmentation process using high and low pass filters, alongside operations to erode/dilate the outcome. The outcome will be regions of connected voxels that in the end may be considered as particles.

In this experiment we have identified 1 239 connected voxel regions at this stage, that is, 1 239 particle candidates.

Fig. 4 shows a network of modules to support the labelling process. Notice that some modules, like those related to pass filters, also output information to visualization modules so specialists can figure out the results of intermediate operations. This includes drawing histograms.

B. Particles detection

At this stage the goal is to figure out the proper boundaries of the real particles. It implies carrying out careful analysis in relation to potential regions of particles that have been considered in the previous stage, so we will end up with particles of potential interest, with proper closed boundaries.

As shown in Fig. 5, particles boundaries are not continuous at the beginning. Therefore, first it is required a reconstruction of the boundaries, which is done using the Poisson surface reconstruction algorithm. Only then we can compute the exact number of particles and respective size.

In this case-study, some of the 1 239 regions of connected voxels originated in the previous stage can still be considered as noise. Therefore, we have used a module to discard those *fake* particles. The cut-off size value set was 100 voxels, which is a value somehow derived from the pre-understanding and knowledge of the specialist about the sample. As result, there were 202 particles with acceptable size, that were then submitted to the Poisson surface reconstruction algorithm. The final outcome was a set of particles, described via a set of *Ply* files and representing polygonal meshes that can be visualized.
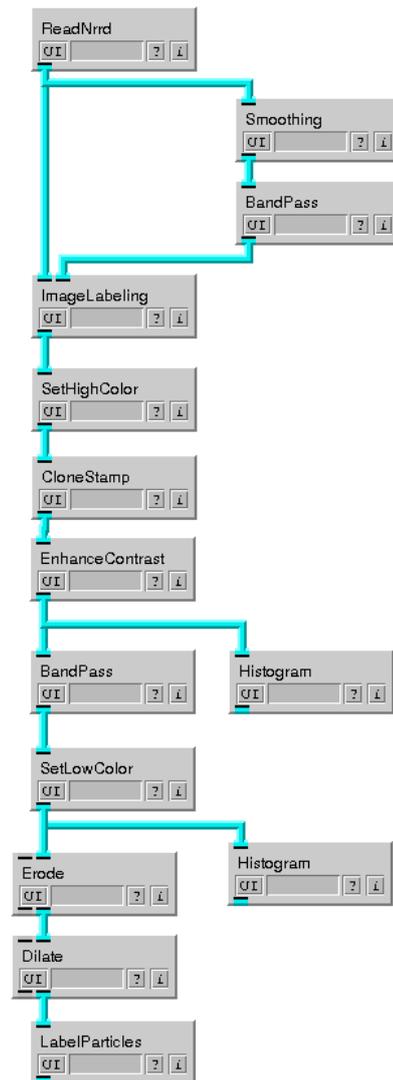

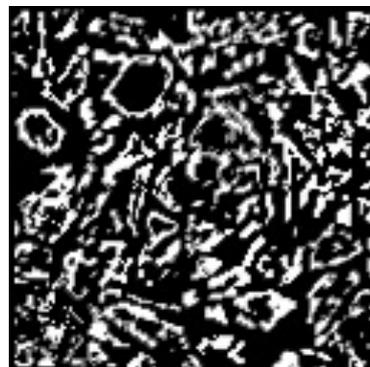Fig. 4. Network of modules set by a specialist with the purpose of supporting particle labelling.

Fig. 5. Initially, particles boundaries are not continuous so proper identification and reconstruction is required.

C. Particles characterization

At this final stage, the goal is to deliver the set of particles the specialist is interested on, and with detailed characterization, mostly based on the geometric profile. It matters not only the selection itself but the quality of characterization.

As depicted in Fig. 6, given the detected particles from the previous stage (Ply files), which may not be entirely correct from a semantic point of view, we enter into an iterative process where, at each iteration, the specialist can accept a particular particle, or reject it, or else submit it to a enhancement process, likewise in the previous stage.

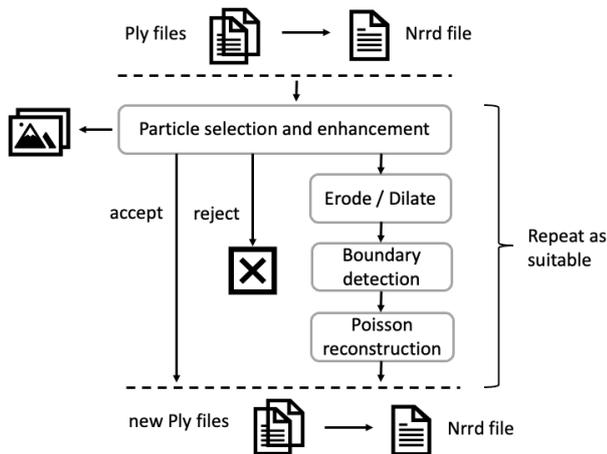The decisions made are also supported by the viewing of particles, as highlighted in Fig. 7.

Fig. 6. Iterative process to select and enhance particles of interest for further characterization.
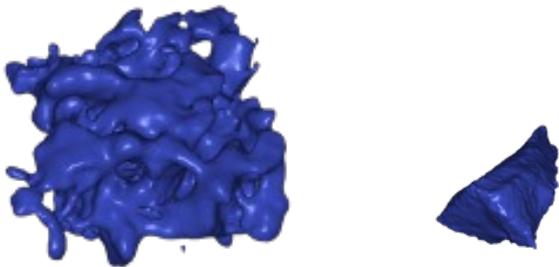

Fig. 7. Visualization to help selecting and enhancing particles of interest: From cluster requiring further processing (left) to accepted particle (right).

Also, the geometric profiles of the particles of interest are computed and stored in a SQLite database. Among other features, it includes the location within the sample, surface area, volume, bounding box, etc. It is worth noting that specialists have general pre-understanding about the samples they are working with, namely in relation to shape and average particle size.

The information stored in the database can be exported to files for further usage with other tools of convenience.

Nonetheless, the implemented *FeatureVisualizer* module is tailor-made for the purpose of visualizing particle features within the context of the PSE itself. The underlying thinking is that it is important to provide extra flexibility to specialists.

VI. Conclusions

We have presented a dedicated PSE to help material science specialists to carry tasks of analysing tomographic images. The proposed solution fits into the list of requirements set by specialists, who were keen on having an environment where (i) they could easily use a wide range of algorithmic strategies to carry out their experiments, (ii) the build up of the processing network was done in a flexible manner and (iii) they should experience human-computer interactivity as much as possible, alongside fast and effective visualizations.

As highlighted in the discussion about particles identification in Section V, tomographic images with low contrast pose additional demands as far as type and number of operations that have to be included in the processing network. In the case-study introduced, the system delivered the results specialists were looking for from a scientific perspective. Also, the network of modules was relatively easy to set up and its steering afterwards was effective.

As a final note, once particles are properly identified and stored in a database, and having various data formats at disposal, a specialist can also use external tools to further analyse the outcome of the experiment, all but in a cohesive working environment.

References

[1] R. Haber and D. McNabb, *Visualization idioms: A conceptual model for scientific visualization systems*. IEEE, 1990, p. 74–93.

[2] D. Stalling, M. Westerhoff, and H.-C. Hege, "Amira – a highly interactive system for visual data analysis," in *The Visualization Handbook*, C. D. Hansen and C. R. Johnson, Eds., 2005, pp. 749–767.

[3] W. Burger and M. Burge, *Digital Image Processing in Java*. Berlin, Heidelberg: Springer-Verlag, 2007. ISBN 1846283795

[4] U. Ayachit, *The ParaView Guide: A Parallel Visualization Application*. Clifton Park, NY, USA: Kitware, Inc., 2015. ISBN 1930934300

[5] O. Stamati, E. Andò, E. Roubin, R. Cailletaud, M. Wiebicke, G. Pinzón, C. Couture, R. Hurley, R. Caulk, D. Caillerie, T. Matsushima, P. Bésuelle, F. Bertoni, T. Arnaud, A. Ortega Laborin, R. Rorato, S. Yue, A. Tengattini, O. Okubadejo, and G. Birmpilis, "spam: Software for practical analysis of materials," *Journal of Open Source Software*, vol. 5, p. 2286, 07 2020. doi: 10.21105/joss.02286

[6] D. M. Weinstein, S. G. Parker, J. Simpson, K. Zimmerman, and G. M. Jones, "Visualization in the scirun problem-solving environment," in *The Visualization Handbook*, C. D. Hansen and C. R. Johnson, Eds. Academic Press/Elsevier, 2005, pp. 615–632. ISBN 978-0-12-387582-2

[7] S. Computing and I. I. (SCI), "Biopse: Problem solving environment for modeling, simulation, image processing, and visualization for biomedical computing applications," 2014. [Online]. Available: http://www.scirun.org

[8] B. Preto, F. Birra, A. Lopes, and P. Medeiros, "Object identification in binary tomographic images using gpgpus," *International Journal of Creative Interfaces and Computer Graphics*, vol. 4, no. 2, p. 40–56, 2013.

[9] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, ser. SGP '06. Goslar, DEU: Eurographics Association, 2006. ISBN 3905673363 p. 61–70.

[10] A. Velhinho, P. Sequeira, F. B. Fernandes, J. D. Botas, and L. A. Rocha, "Al/sicp functionally graded metal-matrix composites produced by centrifugal casting: effect of particle grain size on reinforcement distribution," in *Materials Science Forum*, vol. 423, 2003, pp. 257–262.