Kitware, Inc.

Open Source Visualization and Analysis Platform for 3D Reconstructions of Materials by Transmission Electron Microscopy

Final Report for SBIR Phase IIB Award (Topic 5, Subtopic A)

Period of Performance: April 10, 2017 to April 9, 2020

Due: May 31, 2020

Grantee: Kitware, Inc
1712 Route 9, Suite 300,
Clifton Park, NY 12065

DUNS: 01-092-6207

Principal Investigator: Marcus D. Hanwell
(518) 881-4937
marcus.hanwell@kitware.com

Subawardee: The Regents of the University of Michigan
Principal Investigator: Robert Hovden

DOE Program Manager: Jane Zhu
(301) 903-3811
jane.zhu@science.doe.gov

No data included in this report is marked proprietary

This material is based upon work supported by the U.S. Department of Energy, Chicago Office Under Award Number DE-SC0011385

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Summary

Three-dimensional characterization of materials at the nano- and meso-scale has become possible with transmission and scanning transmission electron microscopes (S/TEM).[1–3] Its importance has extended to a wide class of nanomaterials—such as hydrogen fuel cells,[4] solar cells,[5] block-copolymer networks,[6] low κ-dielectrics,[7] and semiconductor devices,[8] as well as spanning high-tech industry, universities, and national labs. While capable instrumentation is abundant, this rapidly expanding field of S/TEM tomography is bottlenecked by software that is non-existent, limited and proprietary, or tailored for biological applications. The 3D characterization of materials in S/TEM critically relies on software with distinct requirements for data acquisition, alignment and reconstruction, and visualization and analysis that is different from the more established biological tomography. Streamlining this process requires open, extendable software capable of handling all steps.

Kitware and the University of Michigan partnered on the Phase IIB SBIR project to extend an open source S/TEM tomography package for materials with a modern user interface that enables alignment and reconstruction of raw tomography data. It also provides an advanced, efficient 3D visualization and analysis environment for reconstructed volumetric data. It is an extendable framework that has been improved to be useful to more types of tomographic characterization, and to handle more complex S/TEM tomography data involving multiple sensors/fields.

The project has been named '*tomviz*', and combines existing in-house tomography software developed at Cornell University from the Phase I/II project with professional software interfaces developed by Kitware to create a user-friendly, cross-platform, graphical desktop application that offers an integrated collection, alignment, reconstruction, segmentation, and analysis environment. In addition to providing significantly improved, intuitive software, the application enables reproducible workflows where the full path of data from collection to final analyzed results can be saved, shared, and reproduced.

tomviz is available for download at `www.tomviz.org/downloads`. The project had $\sim$ 300 external downloads from August to November 2014 in Phase I growing to over 5,000 downloads by July of 2017 with over 1,800 downloads of the 1.0 release (Phase II). This has now grown to over 15,000 downloads in May 2020 including the latest 1.8.0 release (end of Phase IIB project). It offers a powerful community tool, and demonstrates what is possible using the VTK,[9, 10] ParaView[11] and ITK[12] frameworks coupled with Python in a graphical application designed for materials scientists. It has been developed with large volumes in mind, but tested and optimized to work on commodity laptops typically available to researchers—running on Windows, macOS, and Linux with binary installers offered for all platforms.

Completion of Phase IIB offers real-time tomographic reconstruction, "live updates" as data is processed, and analysis of big data and multi-channel tomography. Here, quantitative assessment of materials can occur as data is being recorded on a 3D visualization platform that accommodates chemical information from multiple techniques. The platform has been integrated at synchrotron beamlines such as NSLS-II, the pipeline is extendable at runtime using Python and JSON, and data has been analyzed using electron, X-ray, and neutron tomography as well as support for FIB-SEM and atomic electron tomography.

The Phase IIB effort greatly enhances commercialization by offering one of a kind real-time tomography that tailors to industry research across many 3D characterization techniques. Building upon the open source, cross-platform Phase II tool this IIB project targets five major opportunities in nanomaterials tomography:

- Real-time automated data acquisition and reconstruction
- Execute reconstruction or processes using remote high-performance supercomputers
- Multi-channel tomography of nanomaterials
- Enhanced pipeline supporting workflows for data collection, analysis, and visualization (one-click recipes)
- Expansion into micro-CT, X-ray, neutron, FIB-SEM, atomic electron tomography and atom probe tomography

These ambitious goals has been be accomplished by extending the carefully constructed Phase II project built as a scalable software platform for state-of-the-art S/TEM tomography of nanomaterials. The Phase IIB project was executed as planned resulting in a powerful open source, cross-platform graphical application focused with hardware accelerated 3D rendering for materials tomography. The tool is well known by the community, has been used to generate figures used on covers, and has resulted in commercialization activities with multiple partners.

Abstract

Three-dimensional characterization of materials at the nano- and meso-scale has become possible with transmission and scanning transmission electron microscopes (S/TEM). Its importance has extended to a wide class of nanomaterials—such as hydrogen fuel cells, solar cells, industrial catalysts, new battery materials and semiconductor devices, as well as spanning high-tech industry, universities, and national labs. While capable instrumentation is abundant, this rapidly expanding demand for high-resolution tomography is bottlenecked by software that is instead tailored for lower-dose, biological applications and not optimized for higher-resolution materials applications. Existing tomography fails to utilize the chemical information provided by S/TEM spectrometers.

To address this problem, this project delivered a scalable, fully functional, freely-distributable, open source materials tomography package with a modern user interface that enables automated acquisition, alignment, and real-time reconstruction of raw tomography data, and provides advanced segmentation, three-dimensional chemical visualization and analysis optimized for materials applications. It has established an extendable framework capable of automation for high-throughput for the tomography of materials from data acquisition to visualization.

Phase I and II delivered a full-featured, cross-platform, clean and integrated application for S/TEM materials tomography. It can read projection data from the microscope, with graphical tools for alignment of data, tomographic reconstruction, segmentation, and visualization of the reconstructed 3D volume. The entire pipeline can be saved to an XML state file, enabling fully reproducible data processing and analysis with a full Python environment for the development of custom algorithms, processing, and analysis requirements.

Phase IIB extended the application to provide real-time tomographic reconstruction, "live updates" as data is processed, analysis of big data, and multi-channel tomography. Here, quantitative assessment of nanomaterials can occur as data is being recorded on a 3D visualization platform that accommodates additional information from multiple channels. We provide unmatched high-throughput tomography, where the complete tomographic pipeline from measurement to 3D visualization can occur rapidly. The project was extended to offer capabilities for micro-CT, X-ray, neutron, focused ion beam, atomic electron tomography and atom probe tomography.

With over 600 transmission electron microscopes worldwide and approximately 50 coming online each year, the demand and impact of an open-source tomography tool is large. Significant opportunities exist in high-tech industry, universities, and national labs to enable or enhance three-dimensional imaging at the nanoscale and bring automated high-throughput approaches that will accelerate progress in materials characterization and metrology. The project supports a service based business model by enabling lab-specific acquisition and processing customization and integration—support and development that will be provided into Phase III and beyond.

Contents

1 Hypotheses **5**

2 Methods/Assumptions/Procedures **6**

3 Results/Findings **7**

3.1 Dissemination to Communities of Interest 14
3.2 Websites/Other Internet Sites 17
3.3 Technologies or Techniques 17
3.4 Impact of the Project and Contributions 18
3.5 Impact on the Development of the Principal Discipline(s) of the Project 18
3.6 Impact on Technology Transfer 19

4 Recommendations **20**

1 Hypotheses

Completion of Phase IIB will provide real-time tomographic reconstruction and analysis of big data and multi-channel tomography. Here, quantitative assessment of nanomaterials can occur as data is being recorded on a 3D visualization platform that accommodates additional information. We will provide unmatched high-throughput tomography, where the complete tomographic pipeline from measurement to 3D visualization can occur rapidly.

The Phase IIB effort enhances commercialization by offering one of a kind real-time tomography that tailors to industry research across many 3D characterization techniques. Building upon the open source, cross-platform Phase II tool this IIB project targets five major opportunities in nanomaterials tomography:

- Real-time high-throughput data acquisition and reconstruction
- Execute reconstruction or processes using remote high-performance supercomputers
- 3D chemical tomography of nanomaterials
- Enhanced pipeline supporting workflows for data collection, analysis, and visualization (one-click recipes)
- Expansion into micro-CT, X-ray, neutron, and atom probe tomography

These ambitious goals will be accomplished by extending the carefully constructed Phase II project built as a scalable software platform for state-of-the-art S/TEM tomography of nanomaterials.

Achieving these high-level goals will require work on a multitude of lower-level components detailed in the following sections. Software development experts will work with experts in S/TEM microscopy to develop this targeted application. This application will be developed collaboratively between Kitware, the University of Michigan, and Cornell University, with regular in-person meetings to discuss needs and capabilities for the 3D characterization of materials. Workshops, tutorials, scientific meetings, and site visits will be used to gain feedback, with regular releases to encourage wider adoption.

The Gantt chart in the proposal showed major milestones/deliverables, addressing the extended range of tomography workflow for a more general audience:

1. Real-time tomographic reconstruction
2. 3D chemical and multidetector tomography
3. One-click recipes for 3D metrology
4. High performance computing for rapid tomography and big data
5. Expansion into other fields
6. Engagement, dissemination, and outreach

Regular software releases will be made to disseminate updated features and gain wider feedback on developments. These will be augmented with automated build and upload of installers from the main development branch to facilitate rapid feedback from engaged community members requesting new features or bug fixes in between releases.

2 Methods/Assumptions/Procedures

Major milestones encompass software releases, with twelve releases made in the reporting period (eight feature releases and four bug fix releases), and another release planned soon with the latest developments that are currently available in the master branch. These releases reflect the most visible milestones, and each feature release contains a number of new features, with bug fix releases offering fixes for bugs discovered after feature releases. In addition to releases, installers are built and automatically uploaded to support greater developer-user interaction as features are developed or bugs are fixed.

The releases made in the Phase IIB project offer significant advances in the areas outlined in the project work plan. At the end of the reporting period the project had significant advances described in the following sections addressing the areas outlined in the previous section. Blog posts for releases summarize the features added, and dissemination efforts highlight them in the community. As this reporting period ended the team has made the 1.8 release which feature more improvements to "live updates" of the pipeline, further reductions in the memory required, improved support for X-ray tomography, and improved use of Docker.

Development is performed openly using GitHub to host distributed version control, issue tracking, continuous integration testing, and distribution of release binaries.

There were also a number of branches that were held back as they weren't quite ready that offer support for executing data processing pipelines in an external Python environment, and for manipulating the data pipeline from the Python environment within the application. These enhancements usher in support for more advanced use cases, such as a native Python that can use GPU-accelerated algorithms, custom Python modules, and advanced scripting of the user interface.

3 Results/Findings

The Tomviz application now enables tomographic reconstruction that updates as **reconstruction** proceeds, in real-time. Scientists are greatly alleviated by the computational bottleneck of reconstruction times, by being able to immediately begin interpreting results, as an algorithm proceeds. This is functional, using simple reconstruction algorithms to offer interactive speeds for small-to-medium sized volumes on commodity hardware. The development of an extended programming interface for the Python reconstruction operators is in place and being used to offer greater interactivity for those algorithms.

The 1.2.1 release was the last made before the pipeline underwent a major rewrite, with 1.3.0 being a significant release where the user interface for the pipeline was enhanced. Beyond the visual aspects, the underlying code for the pipeline is much more flexible, which has the infrastructure necessary to accomplish several the Phase IIB goals. The "new" pipeline code enabled us to move forward with our plan for live updates of the pipeline, adding support for live updates in the user interface as reconstruction proceeds, as well as rerunning the pipeline during data acquisition.

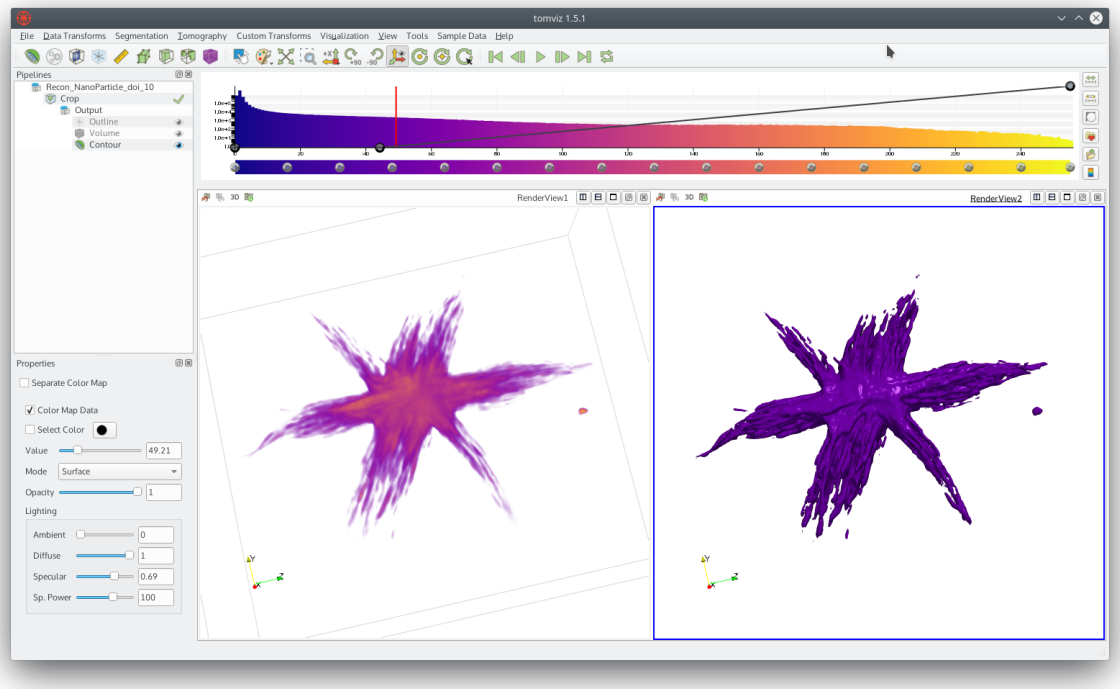

Figure 1: Tomviz displaying a nanoparticle with volume rendering (left) and isocontour (right). The histogram is displayed at the top (and can be collapsed away), with the data processing pipeline and active pipeline object properties on the left.

The rewrite of the pipeline necessitated a rewrite of the state file code to support the increased flexibility of the pipeline, as outlined in the work plan the new state file uses JSON to represent the pipeline in a more concise form. This resulted in an order of magnitude decrease in state file size over the previous XML-based state file through a combination of the more concise form of JSON and greater selectivity of state file items being saved correlating to elements exposed in the user interface. The updated user interface is shown in Figure 1, with the new data pipeline.

Figure 2 shows an example of a snippet of the state file from Figure 1. The JSON format is a simple way of representing objects, arrays, and strings, numbers or Boolean values. The primary data source is shown at the top level, and within that object an array of operators—in this case just one operator, the "Crop" operator type. Its main property is the "bounds" which represent the crop region in pixels. An output of that is the "Output" data source, and this has child modules that act on the output. The "Countour" module has a "contourValue" which represents the value of the surface calculated. This is enough data for

```
"dataSources": [
    {   "operators": [
            {   "bounds": [  0, 99, 0, 215,  0, 250 ],
                "dataSources": [
                    {   "label": "Output",
                        "modules": [
                            {   "properties": {
                                    "gridColor": [  0.8, 0.8, 0.8 ],
                                    "gridLines": false,
                                    "gridVisibility": false,
                                    "visibility": true
                                },
                                "type": "Outline",
                                "viewId": 2200
                            },
                            {
                                "properties": {
                                    "blendingMode": 0,
                                    "interpolation": 1,
                                    "lighting": {
                                        "ambient": 0,
                                        "diffuse": 1,
                                        "enabled": false,
                                        "specular": 1,
                                        "specularPower": 100
                                    },
                                    "rayJittering": true,
                                    "transferMode": 0,
                                    "visibility": true
                                },
                                "type": "Volume",
                                "viewId": 2200
                            },
                            {
                                "properties": {
                                    "activeRepresentation": {
                                        "color": [
                                        ],
                                        "lighting": {
                                        },
                                        "mapScalars": true,
                                        "opacity": 1,
                                        "representation": "Surface"
                                    },
                                    "contourValue": 49.21,
                                    "useSolidColor": false,
                                    "visibility": true
                                },
                                "type": "Contour",
                                "viewId": 3140
                            }
                        ]
                    }
                ],
                "type": "Crop"
            }
        ],
        "reader": {
            "fileNames": [
                "../Recon_NanoParticle_doi_10.1021-nl103400a.tif"
            ],
            "name": "TIFFSeriesReader"
        }
    }
```


Figure 2: Tomviz JSON state file example snippet—not complete but it shows the general form.

Tomviz to recreate the complete pipeline, for brevity the views were omitted but they contain the camera matrix to restore the view. The use of JSON means that it is possible to load this JSON file in Python, manipulate values, and then load the derived state files in Tomviz. The tree displayed in the pipeline maps to the tree in the JSON objects such that motivated researchers can see the relation.

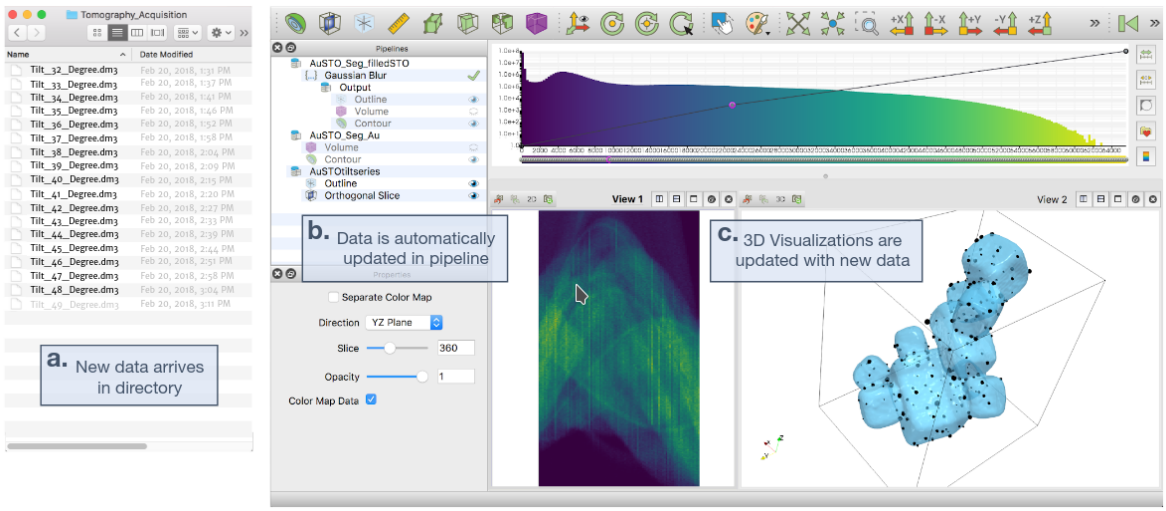

Figure 3: Visual example of Tomviz being used to watch a directory and perform live acquisition workflows on a TEM microscope.

An approach to passive data collection has been designed, and an implementation using our **acquisition server** has been shown working on complete data, with Figure 3 showing an example session. This augments our original approach that uses microscope programming interfaces, enabling data collection in whatever package the microscopist uses. The acquisition process watches the named directory and sends a notification when new images are available. Relevant metadata is extracted, including the tilt angle, to offer real-time interaction with data as it is acquired.

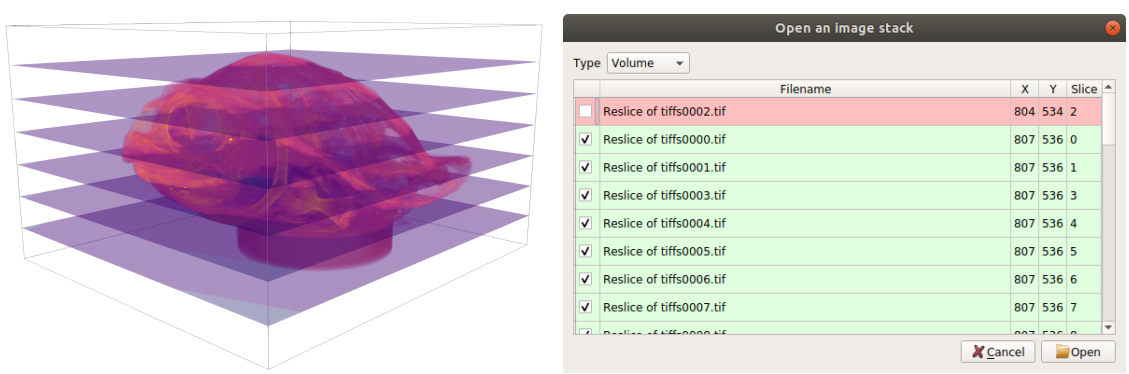

Figure 4: Tomviz features enhanced capabilities to load image stacks such as the one shown here (left) loaded from individual TIFF images (right). The dialog shows a "bad" image with a non-conforming size that will be skipped.

A similar approach was used to add an enhanced image stack dialog where multiple naming schemes can be specified for the stack. User feedback indicated this was an issue for people getting acquired and/or reconstructed data into Tomviz. Regular expressions can be used, with common naming schemes offered where angles or sequence numbers are encoded as a part of the file name. The loading of volumes remains unchanged, this feature serves to augment image stack loading as either a tilt series or as a volume. The dialog displays the size of the images that will be loaded, and refuses to load images with a differing size.

Upon loading the angle data is read in if it is contained in the file name, or it can be entered later. For large numbers of images this will be skipped in order to offer a faster experience when directories are well constructed, but the order of loading will still be displayed so that it can be verified before reading in data, Figure 4 shows the dialog with a non-conforming image.

Support for **chemical and multidetector data** has been added to the Tomviz pipeline, and the team has improved the interface to expose multiple "channels" in a single data set ready for the next release (1.9.0) that will be made soon. There is new user interface to change the active channel for a data source, browse available channels, and features exposing the use of different channels for coloring contours. We have some file formats that natively support multiple channels and have also developed user interface that can combine several files of the same dimensions into a multiple channel data set. The Python interface supports processing multi-channel data, and even outputs with more than one channel.

This support has primarily focused on the visualization pipeline, and on changing the current channel while maintaining a single channel pipeline. More recent development has focused on extending relevant algorithms to use multiple channels from a single data source, and improve data flow through the processing pipeline. This involves a number of updates to the programming interfaces made available within the pipeline, along with interface updates to expose the additional aspects of the data source. The multiple channel support will be fully exposed, but currently is not available in the released version.

This support is still evolving, and despite some delays in completion we have a plan to support multiple channels that is compatible with loading larger data sets. After actively engaging both the TEM community and many members of the X-ray light source/neutron community we have formed a full picture of what is needed in this challenging area. The way in which image stack images with multiple channels need to be aligned is another area of active exploration, with a high signal-to-noise ratio often only available on one channel dominating the alignment. The image stack dialog now supports saving and loading offsets to support some of these workflows in a manual fashion, recent development in master enables switching between loaded channels.

The development of an enhanced pipeline to support **one-click recipes** is in now in place. This is quite a major update, and includes some deep changes that support other aspects of the Phase IIB project. Online update meetings, a tutorial held after the Frontiers of Electron Tomography meeting, a hackathon held at Oak Ridge National Laboratory for X-ray/neutron tomography, and a two day in-person meeting at the University of Michigan were used to develop the updated design, and prioritize updates. These were also used as an opportunity to get direct feedback from our user community, experts, and potential users attending workshops.

The refactoring of the data pipeline has taken this feedback into account, and offers a more flexible pipeline that more accurately reflects what is happening in more complex workflows. Some cosmetic issues were also improved, the output of the operators is now explicitly represented in the pipeline browser, enabling users to visualize the original input data and the processed data more easily. It also supports deeper, more branched pipelines, while retaining the ability to construct simple pipelines with ease.

The updated pipeline required the extension of the state file to support a more general pipeline layout. This focused on using a simpler JSON data container to express the application state. This has the advantage of using simpler parsers that have great availability in many programming languages, and in supporting the HPC task by offering a state format easily used from the Python language. We have been able to develop conventions useful to both the state file serialization/deserialization mechanism and for the packaging of task s for external execution. Figure 2 shows a snippet of a simple pipeline saved to JSON, with some explanation of the object representation used.

The first **HPC** task of packaging pipeline tasks is complete, with working code uses the JSON state file mechanism to package linear pipelines. Local execution in a separate process has been demonstrated, using the Docker container technology to execute pipeline tasks, and offers a completely reproducible pipeline execution framework by using self-contained Linux containers that have a snapshot of all code in the image, downloadable and usable across machines. This goes beyond what has previously been possible where each Tomviz version came with updates to its C++ and Python environment.

The Docker containers themselves are open and can be extended to add new Python packages or other dependencies using industry standard techniques. The desktop application has a setting to switch to the experimental container-based pipeline while we extend it to work with all appropriate operator types. For the supported operators the use of Docker is seamless, with application updates provided by communicating

with the process in the Docker container. The application saves the input volume as an EMD file, and the last step saves en EMD output file using JSON to drive the execution of the pipeline.

We have been working on remote execution strategies, and highly parallel execution of operators, we have pipeline execution capable of running in Docker or a standard Python 3 environment. We continue to explore what can be accomplished in a user friendly desktop application. Pipeline tasks are fully self-contained at this stage, with simple operators run completely independent of the Tomviz application.

An external Python environment can be used, such as a Conda environment commonly employed by many scientists. This is often simpler for end users than setting up Docker containers, and can offer a good path for evolving scripting environments with local packages installed. The Tomviz pipeline runner can also execute pipelines in batch on one or more inputs enabling high throughput execution unsupervised for later visualization/analysis.

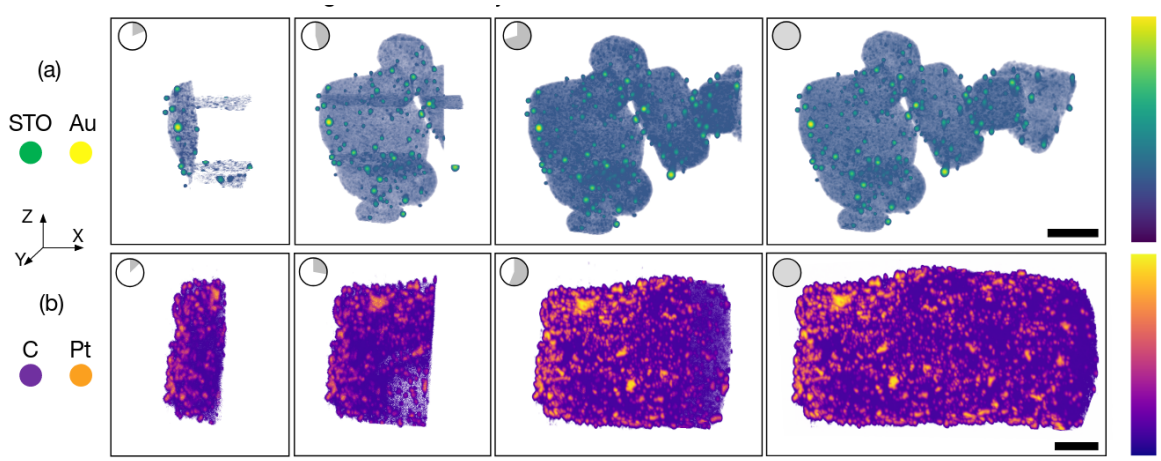

Figure 5: Some snapshots of "live" reconstructions where the application updates throughout the reconstruction process, enabling researchers to view reconstruction as it proceeds.

The reconstruction operators have been significantly improved to reduce memory usage, moving from double precision floating point representations to single precision floating point at most. The algorithms were examined critically to reduce unnecessary creation of large volumes, and where their creation was necessary the lifetime of those objects was reduced where possible. Some operations where sparse arrays were expanded out to dense arrays have been augmented to operate within a more limited scope, this has led to some increase in execution time but it has hugely reduced the overall memory requirements during reconstruction. A hackathon hosted at Kitware in January, 2019 focused on this issue, along with ironing out bugs/interaction issues during live reconstructions, such as those shown in Figure 5. These improvements were initially featured in the 1.6 release, with further improvements in the 1.7 and 1.8 releases.

Live tomographic reconstruction, which allows updated visualization while a algorithm proceeds, has become refined and impactful. Figure 5 is the first time that nanoscale tomography has been reconstructed with live 3D volumetric visualization. Figure 5 shows screen shots taken from a live WBP reconstruction visualized as time proceeds from left to right. Shown on the top row (Fig. 5a) is a tomogram of gold (Au) nanoparticles on a strontium titanate (STO) and on bottom (Fig. 5b) are platinum (Pt) nanoparticles on a carbon (C) support with the rotation axis along the x − direction. In the actual software, the 3D visualizations are continually updated. Most recently, advanced iterative tomographic algorithms also allow live reconstruction. Figure 6 shows how the reconstruction evolves in real-time for the simultaneous iterative reconstruction technique (SIRT)—allowing accurate assessment of convergence and begin immediate interpretation of the specimen.

Expansion to other fields is a major goal for Phase IIB of this project. Dr. Hanwell (PI) obtained funding from the UK EPSRC in partnership with Martin Turner of Manchester University, UK that is in place to foster greater UK-US collaboration by providing travel grants to "research software engineers" as the program terms it. Dr. Hanwell visited the University of Manchester, Daresbury Lab (central facility lab), DIAMOND (central X-ray synchrotron facility) and ISIS (central neutron spallation source facility),

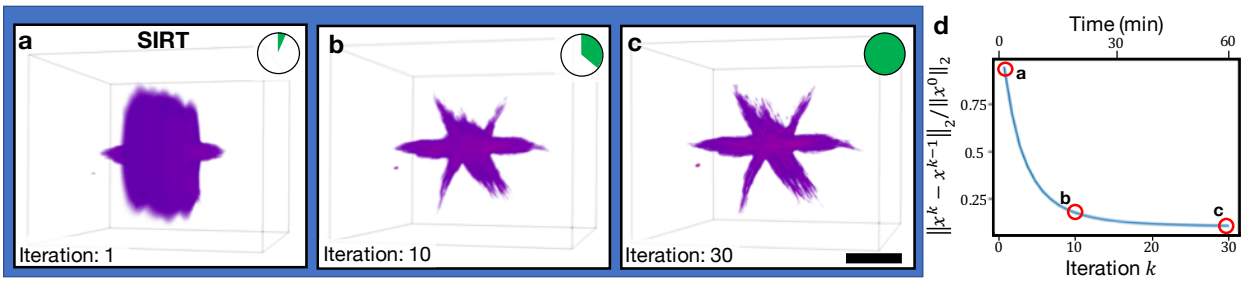

Figure 6: Demonstration of live-SIRT in tomviz. a-c) Visualization of the Co2P nanoparticle early, mid, and the end of the reconstruction process. Initially, the underlying structure is partially seen behind the excess of background intensity. In the middle of the process, sharp features begin to form. The final iteration converges with sharp features. d) Plot of change in reconstruction quantifies convergence. Scale bar, 50 nm.

and ToScA (UK-based tomography conference in Southampton, UK this year) between September 4-13, 2017. He gave a number of presentations detailing Tomviz development, and had meetings with user facility groups on how best to coordinate development effort. Multi-channel support was identified as a prime area of great importance in many types of tomographic data analysis.

Dr. Hanwell was invited to a visualization and analysis hackathon hosted at Oak Ridge National Laboratories, with representatives from Brookhaven National Lab and Argonne National Lab also present. He presented the Tomviz platform, and discussed the visualization and analysis needs of the X-ray and neutron central facilities at major US user facilities. This had a number of other presentations discussing relevant projects, and support for formats used by other groups as well as typical data sizes/dimensionalities were discussed. Some of the ideas discussed have already been implemented.

A two day meeting in December, 2017 at the University of Michigan with Prof. Hovden's group gave the opportunity for the development team to use an aberration corrected TEM. It was also used as an opportunity to look at the micro-CT and atom probe facilities available there. The data requirements were discussed, along with additional facilities in the Tomviz application that could best serve the needs of these communities.

In August, 2018 Dr. Hovden presented recent developments in Tomviz and Dr. Hanwell also attended the Microscopy and Microanalysis in Baltimore, MD. A new release was made to showcase the latest developments, social media was used to promote our attendance. This offered a great opportunity to meet with researchers, collaborators, and some of the microscope vendors to discuss the need for programmatic access to microscopes. The following week Dr. Hanwell attended a hackathon hosted at Brookhaven National Lab to discuss the software needs of X-ray light sources/neutron facilities with strong attendance from US and European-based facilities.

In October, 2018 Kitware sponsored the "NOBUGS" workshop, and Dr. Hanwell presented Tomviz along with data management/workflow platform code to the X-ray light source and neutron user facility representatives. This was also a good opportunity to learn from beamline scientists what their needs were, and where the biggest software gaps were in visualization/analysis pipelines. Dr. Hanwell also presented a talk at the Microscopy and Microanalysis meeting in Portland, OR in August 2019 discussing recent developments and improvements for X-ray/neutron tomography in a track mainly talking about X-ray techniques.

Dr. Hanwell presented Tomviz at the NSLS-II User Meeting at Brookhaven National Lab on May 22, 2019, and is working with NSLS-II to adapt Tomviz for their tomography beamline. A visit is planned for mid-February to go over enhancements in the first year, and discuss possible follow on funding. In October 2019 Dr. Hanwell attended a workshop on imaging at Oak Ridge National Lab where he presented work on Tomviz, and ideas for improving visualization and analysis at X-ray and neutron imaging facilities. Dr. Hanwell was also invited to present a Tomviz tutorial in June 2020 to the UK tomography community at an event being hosted by the Diamond facility.

Support for readers and writers written in Python was added, and demonstrations were written that can read/write MATLAB and NumPy binary format files. These formats are in common use across a number of domains, and the general capability makes it easier for anyone to add a small Python script to read/write

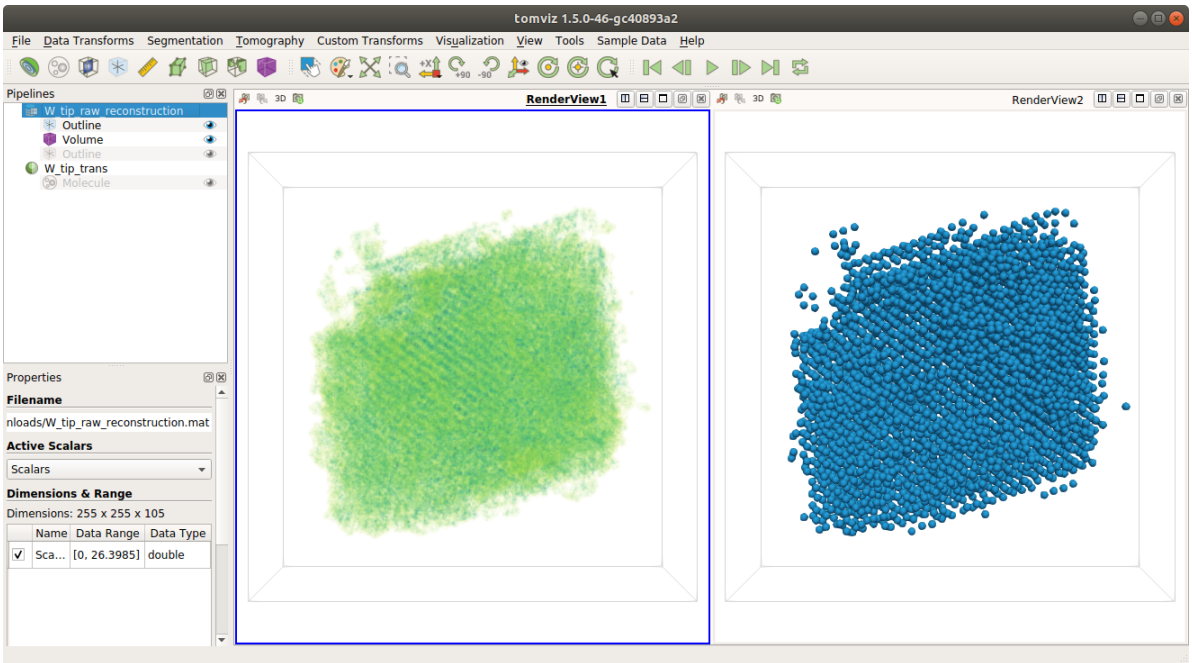

Figure 7: Side-by-side image of atomic electron tomography reconstruction (left) and the atomic data extracted from the reconstruction using atomic electron tomography techniques in the Tomviz application.

to their format of choice. The ability to install custom Python scripts through the user interface was added, offering the capability to place the files in the correct location and add operators for specialized use cases, or prototype new features locally. This feature caused some unexpected side effects, but follow on work has improved initial application start up times, and other issues reported.

Support for 3D sections with FIB/SEM were added, generalizing support for volumes that are not necessarily derived from projection. The alignment tools previously assumed only images with angle data would be aligned, this requirement was removed, and the interface generalized. New algorithms were developed that more effectively remove curtaining by leveraging compressed sensing approaches. This was demonstrated in the Tomviz platform and has been featured in releases. The generalization of image stack loading, and a third volume type has offered improved utility for researchers in these communities.

To better support atomic electron tomography and atom probe tomography initial support for the visualization of atoms has been added and will be featured in the next release. Initially this focused on adding atomic data as a new output type for operators, with a new visualization module that features variable diameter atoms. This will be extended soon to support loading atomic data directly from files, offering views that combine volumetric and atomic visualization in a single application. The atomic coordinates can also be saved to common file formats, useful when detecting atoms in workflows typical of atomic electron tomography.

Figure 7 shows a Tomviz session where the reconstruction of atomic electron tomographic data on the left is compared to the extracted atomic positions on the right. The pipeline in the latest Tomviz releases supports reading in atomic coordinates/elemental data from processed data, or creating atomic position/element data as an output of a pipeline operation. This offers great possibilities for the atom probe tomography community and the atomic electron tomography community to visualize raw data and/or processed results. Additional pipeline features are being examined as our team seeks to serve these communities.

The Python API has been improved in a number of ways, making it more "Pythonic", this involves a number of improvements to make the programming interface look more like other Python-based interfaces. The old programming interface was retained, but no new features were added. The new interface features support for addressing multiple channels in a volume, such as different detectors, and for accessing white/dark fields in X-ray data. Going from the 1.8.0 release we will be recommending new operators are developed

with the new interface, and most of the old operators were converted over.

The support for building data processing pipelines interactively, saving them to a JSON representation and running them in an external Python has been augmented. These pipelines can now be run on many input data files in batch using high-performance computing, clusters, cloud resources and similar. This is being used by the X-ray community to tweak processing pipelines, and then process a series of data acquisitions for later visualization and analysis. The supported data formats have also been extended to encompass those used by electron microscopists, X-ray, and neutron tomography applications, and other related areas.

A "tomvizdocs" repository was also added, and the documentation is automatically built from the latest merged commits in master. They can be seen at https://tomviz.readthedocs.io, and offer extended documentation for the Tomviz application. The site was bootstrapped using the screenshots and examples given in the tutorial slides developed, and have since been extended to include a range of topics. The application has been enhanced with help buttons that link to relevant parts of the online documentation.

The continuous integration used by Tomviz was originally spread across three different services for Windows, macOS, and Linux. The Azure Pipelines offering combines support for continuous integration on all three platforms with a common scripting environment, it also offers greater amounts of time for building software components. This allowed our team to revisit automatically building ParaView when the release used is changed (previously a manual step), caching that build, and then building Tomviz with some of its dependencies in a more consistent way.

Tomviz 1.0.0 was released towards the end of the Phase II project, within the Phase IIB current reporting period a number of software **releases** have been made with full source code, and binary installers for Windows, macOS, and Linux (our minimum goal was to make a software release every six months, and we have generally exceeded that):

- Tomviz 1.0.1—June 14, 2017
- **Tomviz 1.1.0—August 21, 2017**
- **Tomviz 1.2.0—October 22, 2017**
- Tomviz 1.2.1—November 5, 2017
- **Tomviz 1.3.0—May 7, 2018**
- Tomviz 1.3.1—May 14, 2018
- **Tomviz 1.4.0—August 1, 2018**
- **Tomviz 1.5.0—November 9, 2018**
- Tomviz 1.5.1—December 11, 2018
- **Tomviz 1.6.0—April 5, 2019**
- **Tomviz 1.7.0—July 22, 2019**
- **Tomviz 1.8.0—December 27, 2019**

The no-cost extension has enabled us to continue making the best use of our funding to achieve the goals outlined in the Phase IIB proposal.

3.1 Dissemination to Communities of Interest

Engagement, dissemination and outreach were identified as a focus area for Phase IIB. In addition to the meetings mentioned above Dr. Hanwell presented some initial work on the "Materials Data Bank" at the Frontiers in Electron Tomography workshop held at Lawrence Berkeley National Lab in October 2017, and also prepared a half-day Tomviz tutorial that was hosted immediately after the workshop as part of the same meeting. Features were added to Tomviz to support the meeting, and another tutorial on atomic electron tomography featured custom Python processing scripts coupled with Tomviz to demonstrate how atomic resolution electron tomography data can be processed.

Prof. Hovden gave a Tomviz tutorial at the Cornell-hosted summer school in June, 2017. As planned, four Kitware developers visited the University of Michigan in December, and spent time teaching students how to contribute code to Tomviz. Our developers also received some training on how to use the transmission electron microscope there, and had the opportunity to acquire some atomic resolution data. They observed a tomography acquisition session, and used collaborative spaces available to discuss development ideas.

A Tomviz tutorial paper was published in Microscopy Today, and we used visualizations produced using Tomviz on the cover of that article. Our collaborators continue to engage, a talk with given at the Microscopy and Microanalysis in August, 2017, and further outreach is planned including submissions to present at the next Microscopy and Microanalysis conference. More than ten blog posts have been made thus far on the Kitware blog announcing releases, and describing new features. At least three more are planned within the project time frame.

Outreach has been a rewarding way for us to reach diverse communities who could benefit from our software. We aim to educate the public about work in 3D imaging of nanoscale materials, and to actively engage with interested communities by sharing our software. We provide sample datasets ready for immediate visualization using an intuitive graphical interface, and couple that with an introductory dialog for new users to see a volume rendering of a volume the first time the application is opened.

As an open source tool capable of running on most laptops across all major operating systems there is a minimal financial barrier to using Tomviz, and any community analyzing volumetric data can benefit. This extends well beyond researchers in materials science, and we have worked to actively engage researchers in related fields as a core goal in the Phase IIB project. Travel funding was obtained to visit central X-ray synchrotron and neutron spallation sources in the UK performing tomography, tutorials were offered at a Cornell summer school (June 2017), and a Lawrence Berkeley National Lab workshop (September 2017).

More recently Dr. Hanwell has presented at the NSLS-II user meeting at Brookhaven National Laboratory on May 22, 2019. Kitware also offered free courses at their headquarters in Clifton Park, NY in July 2019—including a Tomviz half-day course. This was attended by approximately 25 people from a variety of backgrounds. a follow on course is planned for the summer of 2020. Dr. Hanwell will presented updates on Tomviz at the Microscopy and Microanalysis conference on August 7, 2019. In June, 2019 Prof. Hovden presented at the DOE Advanced Light Source in Argonne National Lab on advanced computation and live reconstruction in the Tomviz.

Sharing our work broadly, under a permissive open source license has empowered interest in science, technology, and the humanities. We have leveraged a number of tools to reach communities outside of materials science, including: press releases, social media, online publications, seminars, tutorials, and mailing lists. Every major release has been accompanied by posts to Twitter, Google+, Reddit, and the Kitware blog. The PI and key personnel have attended conferences/workshops, with presentations that discussed the project and its goals. Additionally we have visited a number of labs, central facilities, and research centers to present results, and show how the project would be useful to their research.

Two journal covers are shown in Figure 8, both created using Tomviz, and distributed to the microscopy community. We have worked with people in the community to promote the use of Tomviz, and showcase the quality of visualization that can be produced. We have been extremely pleased to get a few cover images, and will continue to pursue further opportunities.

Publications:

1. Tutorial on the Visualization of Volumetric Data Using tomviz, B. D.A. Levin, Y. Jiang, E. Padgett, S. Waldon, C. Quammen, C. Harris, U. Ayachit, M. Hanwell, P. Ercius, D.A. Muller, R. Hovden, Microscopy Today, 26 (2018)

2. Removing Stripes, Scratches, and Curtaining with Non-Recoverable Compressed Sensing, J. Schwartz, Y. Jiang, Y. Wang, A. Aiello, P. Bhattacharya, H. Yuan, Z. Mi, N. Bassim, R. Hovden, Microscopy and Microanalysis, Microscopy and Microanalysis 25 (2019) arXiv:1901.08001

3. A Simple Preparation Method for Full-Range Electron Tomography of Nanoparticles and Fine Powders, E. Padgett, R. Hovden, J.C. DaSilva, B. D.A. Levin, J.L. Grazul, T. Hanrath, D.A. Muller, Microscopy and Microanalysis 23, 1150 (2017)

4. Sampling limits for electron tomography with sparsity-exploiting reconstructions, Y. Jiang, E. Padgett, R. Hovden, D.A. Muller, Ultramicroscopy 186, 94 (2018)

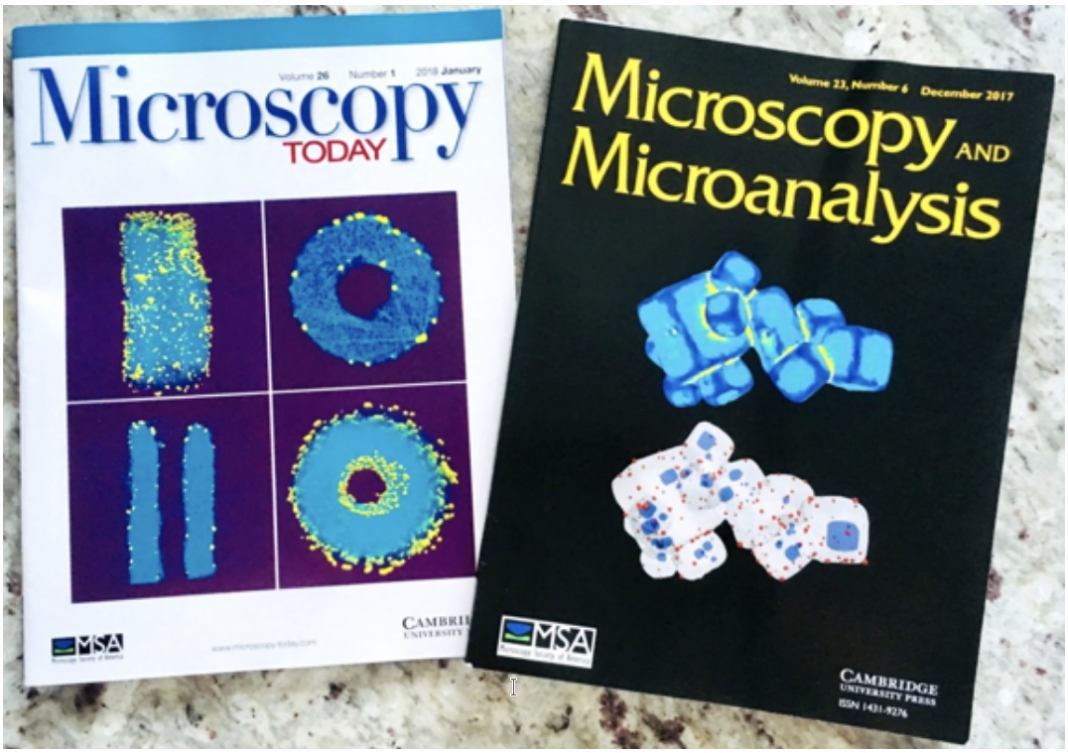

Figure 8: Covers of two recent publications rendered using Tomviz.

Conference Proceedings:

1. Tomviz: Open Source Platform Connecting Image Processing Pipelines to GPU Accelerated 3D Visualization, Marcus D. Hanwell, Christopher J. Harris, Alessandro Genova, Jonathan Schwartz, Yi Jiang, and Robert Hovden, Platform Presentation (Summer 2019)

2. Removing Stripes, Scratches, and FIB Curtaining with Non-Recoverable Compressed Sensing, J. Schwartz, Y. Jiang, Y. Wang, A. Aiello, P. Bhattacharya, H. Yuan, Z. Mi, N. Bassim, R. Hovden, Microcopy and Microanalysis, Platform Presentation (Summer 2019)

3. Real-Time Tomography with Interactive 3D Visualization using tomviz, Robert Hovden, Jonathan Schwartz, Chris Harris, Cory Quammen, Shawn Waldon, Peter Ercius, Marcus D. Hanwell, Yi Jiang, Microscopy & Microanalysis, Platform Presentation (Summer 2018)

4. From 2D to 3D with Reproducible, High-Throughput Electron Tomography, Robert Hovden, Yi Jiang, Jonathan Schwartz, Chris Harris, Cory Quammen, Shawn Waldon, Peter Ercius, Marcus D. Hanwell, Midwest Imaging and Micro-analysis Workshop, Notre Dame (Spring 2018)

5. Removing Stripes, Scratches, and FIB Curtaining with Non-Recoverable Compressed Sensing, J. Schwartz, Y. Jiang, R. Hovden, Canadian Centre for Electron Microscopy, McMaster University (Summer 2018)

6. Removing Artifacts in Micrographs with Compressed Sensing Jonathan Schwartz, Yi Jiang, Robert Hovden, McMaster Focused Ion Beam Workshop, Platform Presentation (Summer 2018)

Workshop Contributions:

1. Cornell Electron Microscopy Workshop, Tomography Lecture + Tutorial (Summer 2017)

2. Lawrence Berkeley National Lab Tomography Workshop, Lecture + Tutorial (Summer 2017)

3. McMaster Tomography Workshop, Tomography Lecture + Tutorial (Summer 2017)
4. McMaster FIB Workshop, Short Seminar (Summer 2018)
5. U. Michigan MAS EBSD Workshop, Keynote Seminar (Summer 2018)
6. Brookhaven National Lab NOBUGS Workshop, Short Seminar (Fall 2018)
7. National Synchrotron Light Source II User Meeting, Short Seminar (Summer 2019)
8. Cornell Electron Microscopy Workshop, Tomography Lecture + Tutorial (Upcoming, Summer 2020)

3.2 Websites/Other Internet Sites

The main site is hosted at tomviz.org with tomviz.org/downloads offering the latest releases, nightly builds with the latest updates, and access to all previous releases via GitHub's release pages. Some focused pages were added such as tomviz.org/publications to highlight publications relevant to the project. Development is hosted on GitHub, github.com/openchemistry/tomviz which features source code, releases, issues, pull requests and other details. A screen recording is on the main page showing some of the features in Tomviz 1.3. A new documentation site is available on tomviz.readthedocs.io, automatically built from a GitHub repository whenever new documentation is added.

Blog posts made through the reporting period detailing releases are also available. All Tomviz related posts can be found at `https://blog.kitware.com/tag/tomviz/`. A summary:

- Tomviz 1.0.1 Released
- Tomviz 1.1.0 Released
- 2D Transfer Function Support in GPUVolumeMapper
- Tomviz 1.2 Released
- CMake Superbuilds and Git Submodules
- Tomviz 1.3 Released: Pipelines, State and Live Updates
- Tomviz 1.3.1 Released
- VTK's Image Comparison Tests and Image Shifts
- Tomviz 1.4 Released: Microscopy and Microanalysis
- Tomviz 1.5 Released
- Opening TIFF Image Stacks in Tomviz
- Tomviz 1.5.1 Released
- Atomic Structures in Tomviz
- Tomviz 1.6 Released
- Tomviz 1.7 Released
- Tomviz 1.8 Released

3.3 Technologies or Techniques

3D rendering, C++, Python, OpenGL, CMake, Qt, Git, visualization, data analysis, materials tomography in a cross-platform, open source project with regular software releases.

3.4 Impact of the Project and Contributions

The impact of the project has been significant, with cover images featuring visualizations from the project, a number of publications using Tomviz in the preparation of manuscripts. Specific impacts are described in the following sections. The broadening of our focus to other types of tomography has enabled our project to have a wider impact which is already leading to new contracts at Kitware. The focus remains on materials, but there are members of other communities attempting to make use of Tomviz from cells through to honey bees! We are also seeing use on much larger samples, with applications from atomic resolution tomography through to macroscale objects such as turbines using neutron tomography.

3.5 Impact on the Development of the Principal Discipline(s) of the Project

Enhanced, streamlined S/TEM tomography for 3D characterization of materials will offer significant public benefit. With around 600 instruments worldwide, and approximately 50 coming online each year, the demand and impact of an open source tomography tool would be large. In Phase IIB we are multiplying this impact by expanding to the wider field of 3D materials characterization that includes micro-CT, synchrotron facilities, focused ion beam, atomic electron tomography and atom probe tomography.

The work will best suit user facilities by providing an open, cross-platform tomography analysis package. For user facilities, educational institutes, or any other setting where multiple users on a single tool can be expected, the limited availability of software becomes a large bottleneck to data analysis, user training, and throughput. For example, Cornell and U. Michigan are each faced with several hundred users logging over 10,000 hours on the instruments and demands a universal data analysis tool that can be freely distributed, and runs on all computers. Commercial software can be prohibitively expensive, and locked to a single machine, requiring users to return to the facilities to simply process or analyze their data. We will provide a solution for central user facilities to not only offer an integrated S/TEM tomography application on-site for materials characterization, but to offer this same application to users for off-site reconstruction and analysis.

The industrial potential is also significant, including large segments of the automotive industry in fuel cell research, and alternative energy solar cell research. In particular, the semiconductor industry's need and spending for nanoscale metrology has continually grown over the last ten years. The International Technology Roadmap for Semiconductors (ITRS) specifically emphasizes the need for 3D characterization of transistors, memory, and interconnects. The next decade will see an increased move toward stacked 3D device structures, such as the FinFET transistor, with feature sizes reaching levels where atomic precision is essential to understand device failure and improve manufacturing processes.

Industry will greatly benefit from the real-time and chemical tomography that will be provided. Streamlining the reconstruction and visualization process will alleviate the bottleneck that prevents the rapid metrology needed by industry. With a platform for complete tomographic analysis developed in Phase IIB, the provision of consulting services will create a market for specialized workflow tools that are largely automated for the typical routine analyses. The flexibility of the tool will also enable research and more challenging analyses within the same environment. Integration into larger workflows, and data management tools will be assets in some settings, along with support services for these facilities.

The requirements for materials tomography are not typically limited by the microscope facilities but rather requires a hard-to-find knowledge-base and software tools. The proposed work will dramatically reduce the barrier of entry for materials tomography in research labs and universities. It will provide well documented software with user-friendly graphical applications developed for the field. Our goal is to develop software that can be run with minimal instruction or training—this means ensuring all options needed for basic analysis are immediately visible to and comprehensible by the user. For researching/developing novel tomographic alignment and reconstruction techniques, these groups will be able to move beyond publishing one-off results and easily integrate their approaches with our platform in order to reach less technically savvy users.

The DOE 2013 Data Crosscutting Requirements Review strongly emphasizes open science as improving the replicability and sharing of data and visualization. For electron tomography, we will develop an open approach where the full pipeline of data acquisition, alignment, reconstruction, segmentation, and analysis of the final data can be presented, saved, and restored. Through the use of state files the application will offer the capability to unambiguously save the workflow pipeline to a file referencing all input data. By enabling fully reproducible results, this has significant implications for work in the area where users will

share data with colleagues, or raise queries with experts. These state files can also be used in the publication of results, where open scientific methods and the additional benefits of wider review can be applied to S/TEM materials characterization. Making use of the web components in the underlying frameworks the same state files coupled with the raw data can also be shared over the web—going from sharing static images of several poses to sharing an aligned, reconstructed 3D volume that can be shown on the desktop, web, and mobile phone. This could be employed by central facilities to highlight important results and to share large data with users after their work was completed.

Tomviz was presented at the Nature's 2016 Scientific Data meeting on data sharing and reuse. The unique architecture, focus on reproducible, editable pipelines and generalization to support a wider array of tomography techniques all improve the outlook for the Tomviz project.

3.6 Impact on Technology Transfer

The importance of electron tomography has extended to a wide class of nanomaterials—such as hydrogen fuel cells, solar cells, block-copolymer networks, low-k dielectrics, and semiconductor devices, as well as spanning high- tech industry, universities, and national labs. While capable instrumentation is abundant, this rapidly expanding field is bottlenecked by software that is non-existent, limited and proprietary, or tailored for biological applications. The 3D characterization of materials in S/TEM critically relies on software with distinct requirements for data acquisition, alignment and reconstruction, and visualization and analysis that is different from the more established biological tomography. We have streamlined this process with an open, extendable software platform capable of handling all steps.

The development in the thirty six months of the project has enabled wider adoption of the materials tomography technique and provides transferable data analysis. Our software tool for S/TEM tomography reconstruction is currently being used by material scientists. To name a subset of institutions that have made their adoption of our tool apparent: DOE facilities Lawrence Berkeley and Brookhaven National Laboratory, Brookhaven National Laboratory, University of Illinois Urbana-Champaign, UCLA, Cornell University, General Motors, and NION Inc.

Invitations and attendance at various events, including two tutorials offered at user-centric microscopy meetings, a hackathon for data visualization/processing of neutron/X-ray tomography data hosted by Oak Ridge National Lab, and meetings at various UK facilities have highlighted wide applicability. These discussions have opened up further possibilities of direct contributions of technology being developed at these facilities/institutions to the Tomviz project, increasing packaged capabilities, transferring more technology from research to the wider community, and forming lasting collaborations. As integration with HPC, Docker containers, and related technologies proceed this situation has improved further, enabling wider adoption of the data pipeline technology being developed in Tomviz.

A contract was in place with UCLA to create a tailored solution for sharing atomic electron tomography data on the web, ending in December of 2018. A contract is in place with Lawrence Berkeley National Laboratory to develop image processing code and visualization capabilities for their new high data rate 4D STEM detector, with an extension/increased funding currently now contracted. Brookhaven National Laboratory have a contract in place that began February 1, 2019 to enhance Tomviz to serve their tomography beam lines at NSLS-II.

A commercial contract was executed for a microscope vendor to receive training on Tomviz in November, 2018, they have issued follow up contracts to develop custom software/consulting to develop visualization and analysis capabilities in their own microscope software tool. Another contract was executed with a semiconductor company to provide on-site Tomviz training in February, 2019. These represent significant milestones in technology transfer, with other projects in the early stages of discussion. A new project funded by the Air Force Research Lab has led to technology transfer and applications in high energy X-ray diffraction, adapting a number of approaches used in Tomviz for this new area of materials characterization and leading to at least $300,000 of investment, with more funding expected.

4 Recommendations

The *tomviz* project has established itself as a vibrant and successful open source project realizing the goals of open science for materials tomography. It has seen uptake in academia, national labs, industry, and government. Beyond the project it has also led to fruitful collaborations on improved data and metadata representations, along with requests for significant expansion into emerging fields of research. We recommend that the DOE continues to fund the development of open software to work hand-in-hand with the latest equipment acquisitions. It is important to push equipment vendors to offer open programming interfaces to control their equipment, and open data standards to capture output.

We recommend that professional software development become more widespread in close collaboration with academics, researchers, and others in order to foster powerful software platforms to enable research. As sensors improve, data capture rates improve, and the complexity of the data increases it is imperative that we develop solutions that can utilize high-performance computing, and leverage the latest advances in artificial intelligence to automate what is currently done manually.

References

[1] De Rosier, D. and Klug, A. Reconstruction of three dimensional structures from electron micrographs. *Nature*, **217**(5124), (1968) 130–134.

[2] Midgley, P. A., Weyland, M., Thomas, J. M. *et al.* Z-Contrast tomography: a technique in three-dimensional nanostructural analysis based on Rutherford scattering. *Chemical Communications*, **10**, (2001) 907–908.

[3] Midgley, P. and Weyland, M. 3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography. *Ultramicroscopy*, **96**(3), (2003) 413–431.

[4] Yu, Y., Xin, H. L., Hovden, R. *et al.* Three-dimensional tracking and visualization of hundreds of Pt-Co fuel cell nanocatalysts during electrochemical aging. *Nano letters*, **12**(9), (2012) 4417–4423.

[5] Kourkoutis, L. F., Hao, X., Huang, S. *et al.* Three-dimensional imaging for precise structural control of Si quantum dot networks for all-Si solar cells. *Nanoscale*, **5**(16), (2013) 7499–7504.

[6] Sai, H., Tan, K. W., Hur, K. *et al.* Hierarchical Porous Polymer Scaffolds from Block Copolymers. *Science*, **341**(6145), (2013) 530–534.

[7] Xin, H. L., Ercius, P., Hughes, K. J. *et al.* Three-dimensional imaging of pore structures inside low-κ dielectrics. *Applied Physics Letters*, **96**(22), (2010) 223 108–223 108.

[8] Ercius, P., Gignac, L. M., Hu, C. *et al.* Three-dimensional measurement of line edge roughness in copper wires using electron tomography. *Microscopy and Microanalysis*, **15**(3), (2009) 244.

[9] VTK. Online (2020). http://www.vtk.org.

[10] Schroeder, W., Martin, K. and Lorensen, B. *An Object Oriented Approach to 3D Graphics*. Kitware, Inc., 4th edition (2004).

[11] ParaView. Online (2020). http://www.paraview.org.

[12] ITK. Online (July 2020). http://www.itk.org.