

RadVolViz: An Information Display-Inspired Transfer Function Editor for Multivariate Volume Visualization

Ayush Kumar, Xinyu Zhang, Huolin L. Xin, Hanfei Yan, Xiaojing Huang, Wei Xu,
and Klaus Mueller, *Senior Member, IEEE*

Abstract—In volume visualization transfer functions are widely used for mapping voxel properties to color and opacity. Typically, volume density data are scalars which require simple 1D transfer functions to achieve this mapping. If the volume densities are vectors of three channels, one can straightforwardly map each channel to either red, green or blue, which requires a trivial extension of the 1D transfer function editor. We devise a new method that applies to volume data with more than three channels. These types of data often arise in scientific scanning applications, where the data are separated into spectral bands or chemical elements. Our method expands on prior work in which a multivariate information display, RadViz, was fused with a radial color map, in order to visualize multi-band 2D images. In this work, we extend this joint interface to blended volume rendering. The information display allows users to recognize the presence and value distribution of the multivariate voxels and the joint volume rendering display visualizes their spatial distribution. We design a set of operators and lenses that allow users to interactively control the mapping of the multivariate voxels to opacity and color. This enables users to isolate or emphasize volumetric structures with desired multivariate properties. Furthermore, it turns out that our method also enables more insightful displays even for RGB data. We demonstrate our method with three datasets obtained from spectral electron microscopy, high energy X-ray scanning, and atmospheric science.

Index Terms—Volume rendering, volume visualization, transfer function, battery, multivariate data, multi channel data, color mapping.



1 INTRODUCTION

VOLUME rendering is a widely used technique for the visualization of 3D scalar fields in form of 2D projections. Invented more than 20 years ago, transfer functions have been an integral part of volume rendering [1], [2]. They are used to map the volume data to color and opacity in order to emphasize important structures and suppress unimportant ones, affording users a better understanding of the data. Clearly, the ability to design effective transfer functions is crucially important and has been the subject of much research.

Most transfer function design techniques assume that the volume data are scalar quantities, and so the visual properties opacity and color are only driven by a single variable. Kniss et al. [3] added the gradient magnitude calculated from the data and showed how a two-dimensional transfer function augmented by a set of manipulation widgets can better delineate intricate surfaces of volumetric structures. Yet, this technique operates only in the 2D space spanned by the scalar data and their gradient magnitude.

We advance a fundamentally new technique by which users can design transfer functions in data spaces that span more than two dimensions. These types of data occur in applications that generate data from multiple modalities or spectral scanning. Our driving application resides in

material science but our method can also be applied in other fields and applications that produce data where volume elements, or *voxels*, have multiple spatially co-located values, such as multi-modal medical imaging or scientific simulations. All of these applications typically produce several volumetric datasets with complementary information, e.g., the percentage of chemical elements present at each voxel. The goal of our research is to allow scientists to visualize these data together and so directly reveal the composition and interaction of the materials resident at the voxels.

The classical workflow of most scientists in these application areas is to visualize each volume dataset, or *channel*, individually and then slowly build up a composite image. This sequential process makes the analysis process rather tedious and labor-intensive. While mapping three material or modality channels to RGB could potentially speed up this process, there are often more than three such channels that make it infeasible to go this route in general.

To address these shortcomings we devised *RadVolViz*, a visual analytics system to help users visualize multivariate volumetric data in a direct and holistic manner. This is a difficult problem from two perspectives: (1) each pixel in a volume rendered image is an aggregation of volume samples along a rendering ray and so forms a composite of information, and (2) each volume sample itself is a composite of multiple volume channels. To help users untangle this mix of information we provide several displays, and interactions defined on them, that allow users to first gain an overview of the data and then drill into details autonomously. Shown in Figure 1, *RadVolViz* has two fully linked displays: (1) an information display that allows insight on the multivariate

- A. Kumar, X. Zhang, K. Mueller are with the Department of Computer Science, Stony Brook University, Stony Brook, NY. W. Xu, H. Yan, X. Huang are with Brookhaven National Lab, Upton, NY. H. L. Xin is with the Department of Physics and Astronomy, UC Irvine, Irvine, CA
E-mail: {aykumar, zhang146, mueller}@cs.stonybrook.edu, {xuw, huan, xjhuang}@bnl.gov, huolin.xin@uci.edu

Manuscript received xxx; revised xxx.

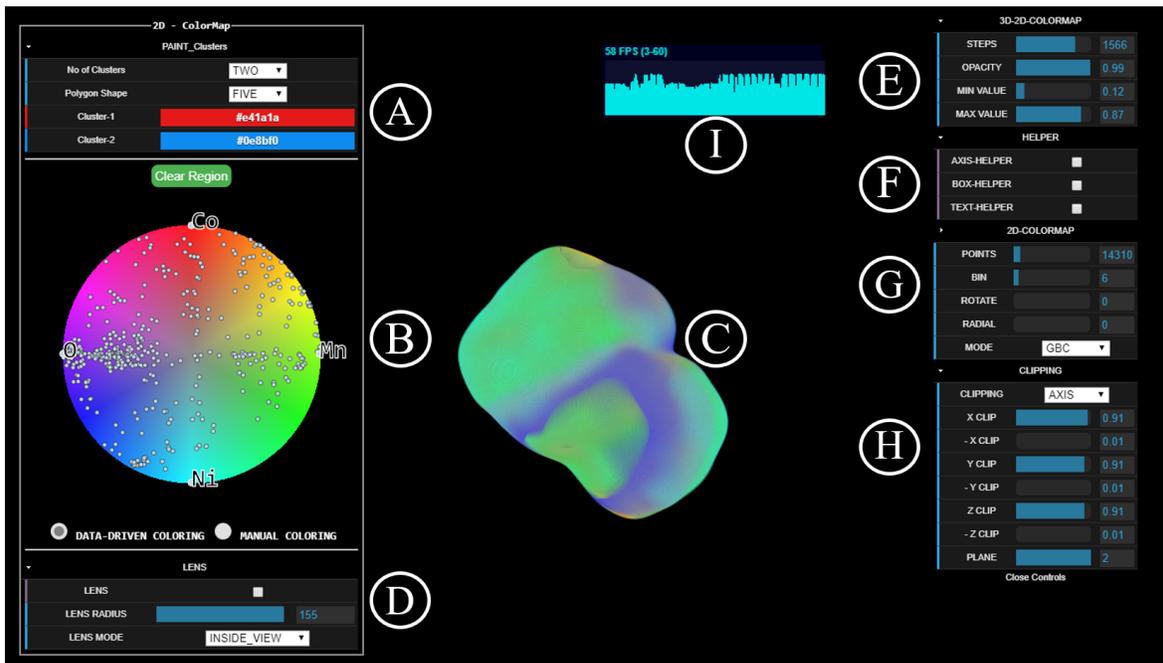


Fig. 1. RadVolViz system interface with all major displays and functionalities, here using the Chemically Sensitive Electron Tomography Dataset. (C) is the rendered volume dataset; (B) is the interactive multivariate information display which doubles as the volume transfer function editor (MTE for short); (A) is used for manual coloring within the MTE; (D) controls the MTE’s lens operations; (E)-(G) are various control elements to tune rendering parameters; (H) is the interface for volume clipping; (I) is a frames/s rendering performance gauge.

aspects of the data and (2) a volume rendering display that visualizes the spatial aspects. Several interactive functions allow users to navigate and explore data aspects of interest.

Our paper is organized as follows. Section 2 gives background on our driving domain application. Section 3 discusses related work on transfer function design. Section 4 presents our design goals. Section 5 describes our approach and implementation. Section 6 presents a set of use cases. Section 7 compares our approach with other modalities. Section 8 presents an evaluation. Section 9 provides a discussion. Section 10 offers conclusions.

2 DOMAIN BACKGROUND AND MOTIVATION

Our domain application is in material science, in particular, to support the quest for more efficient and safer batteries. Batteries are a key driver of many technologies, such as smartphones, electric cars, data centers, and so on. But making strides in battery technology takes an enormous amount of experimental research which in turn requires imaging and subsequently, visual analytics.

Lithium-ion is still considered to be one of the lightest and most efficient battery solutions. It enables energy densities proportional to a cell operating voltage and storage capacity [4]. A higher operating voltage requires cathode materials that are of larger reversible capacity and can withstand higher charging voltages. However, batteries operating at higher voltages undergo various harmful phase transformations and interfacial reactions [5], which raise safety concerns and degrade battery life [6] as well.

The structural stability at higher operating voltages can be improved by elemental doping [7]. Scientists analyze the dopant’s spatial distributions to (1) understand the exact role the dopant plays in stabilizing the chemical environment and (2) optimize the design of the battery’s doping

composition. Especially important is to gain a full understanding of the compositional and chemical heterogeneity that is evolving both on the battery surface and within its material. To obtain this insight scientists perform in-depth qualitative and quantitative evaluations on the doping particles during the electrochemical changes.

Since the electrochemical interfaces at the surface completely change the chemical properties associated with the battery, the study of structural compositions on material surfaces has attracted tremendous research efforts from material scientists. To study the emergent material phases and to quantify the chemical complexity that has occurred due to interfacial reactions at higher voltages, scientists use multimodal hard X-ray imaging at high spatial resolutions [5]. These scans then form the basis for the study of the generated material phases and the complex chemical compositions at a finer scale. It is especially the emerging interfacial reactions that lead to the formation of additional material phases at the grain boundary; these play an important role in determining the ultimate electronic and ionic transport properties of the battery material [8].

Besides the slow data acquisition speed, it is the complexity of the data sets that makes data analysis, visualization, and information extraction a daunting task. Hovden et al. [9] discuss a STEM tomography method that provides quantifiable internal morphology and spectroscopic detection of elements using 3D visualization tools [10], but this method does not scale beyond three channels; analyzing more than three channels is tedious and labor-intensive.

In fact, scientists not only require a visualization tool that can render a multivariate data set to highlight subtle changes in physical and chemical quantities, but they also need a method that allows them to perform the analysis in an autonomous manner. Having these capabilities available

would significantly reduce the time required for analyzing the imaging data, and potentially allow for new discoveries in the evolution of a wide range of materials [11].

These limitations have motivated us to develop a method that can (1) visualize all channels simultaneously to allow scientists to examine the structural composition holistically in one view, and (2) perform the analysis in a semi-automated way to significantly reduce the time to insight.

3 RELATED WORK

The topic of transfer function is one of the most heavily studied subjects in volume rendering [12], [13]. They can be 1D (scalar), 2D, or multi-D, similar to the associated volumetric data. 1D transfer functions are most commonly used, especially in commercially and publicly available software such as ParaView [14], Tomviz [10], and ImageVis3D [15]. When the volume data have just one channel then a simple transfer function interface can assign color and opacity to scalar data ranges, most often via a set of curves. If the volume densities are vectors of two or three channels, one can straightforwardly map each channel to either R, G, or B, which requires a trivial extension of the 1D transfer function editor. The complexity increases, however, when the data dimension of the volume density exceeds three.

To address this complexity, Maciejewski et al. [16] propose an automatic generation of transfer functions using 2D clustering in feature space; temporal variations are encoded in a histogram volume. Tzeng and Ma [17] devise a self-organizing data analysis technique to find clusters for which users can then select color and opacity values. Wang et al. [18] use hierarchical clustering for a similar goal.

Others use dimension reduction to first map the high-D data to lower-D and then design the transfer functions in this reduced space. Kniss et al. [19] achieve dimension reduction through a graph-based technique. Haidacher et al. [20] reduce multi-D data into a single fused representation which they then map through a gradient magnitude two-D transfer function. Kim et al. [21] use a mapping of gradient magnitude into a two-D transfer function using Isomap and PCA as a dimension reduction technique. Isomap dimension reduction is also used by Abbasloo et al. [22].

While these 2D space embeddings can reveal the relations between the volumetric data samples, they lose their associations with the channels the dimensions represent. One solution is that by Mörth et al. [23] who encode the local multivariate data, or parameters derived from them, as a glyph, such as a radar plot, and display them on a volume slice. Users can then appreciate the homo/heterogeneities of different volume regions by comparing the similarities in the field of glyphs. While the glyphs directly encode the local values of the multivariate volume, their spatial extent reduces the spatial resolution. In contrast, colorization can retain the full resolution but an extra data display is needed to show the multivariate data values and to drive the colorization. Guo et al. [24] use parallel coordinates (PCs) to show the data distributions of each channel and MDS to show the dimension-reduced data, all integrated seamlessly for flexible feature classification, while Wang et al. [25] use PCs in conjunction with a multi-resolution

decomposition of the spatial domain. However, PCs have certain shortcomings, as has been pointed out by Bertini et al. [26]. While PCs can show quantitative information well, they tend to clutter and are tedious to manipulate.

We have compared our approach with a simplified version of Wang et al. in Section 7. Apart from the different information visualization interface the approach by Guo et al. also differs from ours in that volumetric structures are first identified via system-assisted visual inspection in (grey-scale) information space, then colorization is applied, and finally, the volume is rendered. Conversely, in our system colorization is immediately and automatically applied and the volume is subsequently rendered. It allows users to recognize clusters in information and in volume space at the same time. In this way, they gain a good idea of what the structures look like in the dual space, and only then they need to delineate and colorize more defined structures in information space and render them in volume space.

Bertini et al. suggest RadViz [27] as an interface with better data manipulation capabilities via brushing (and less clutter), but trading this off with weaknesses in quantitative data display. We chose a RadViz-type information display since we preferred its brushing capabilities while alleviating the less specific quantitative information with visual feedback in the spatial display. We also conducted experiments that use parallel coordinates to assign colors to volume samples but found that the color mapping led to a reduced differentiation of important volume structures. Section 7 provides some of these experimental results.

3.1 Automation

Since the design of transfer functions is a time-consuming and tedious task, several data-driven approaches have been proposed to generate either fully or semi-automated transfer functions requiring little to no user interactions. In this vein, Pfaffelmoser et al. [28] use the spatial deviation of surface points from the mean surface to color the surface, Lundström et al. [29] segment multi-dimensional transfer functions by measures derived from intensity values or spatial relations, Wang et al. [30] let users modify an importance function conveyed via the saliency of the color, Bramon et al. [31] employ notions of informativeness and informational divergence, Sereda et al. [32] apply explorable hierarchical clusterings, Maciejewski et al. [16] use 2D clustering of values and gradients, Ip et al. [33] use a hierarchical exploration of histograms to choose segments hierarchically, and Liu et al. [34] project high-D data to lower dimensions to guide the user through variables; multiple projections are then used to design the transfer functions. Finally, Tzeng et al. [35] and De Moura Pinto et al. [36] achieve dimension reduction through unsupervised learning and then assign voxel properties in the resulting low-D space.

3.2 Coloring

An important aspect of transfer function design has been to shade the volume data with realistic looking colors, in order to emphasize important parts and isolate less important structural portions. Muraki et al. [37] used realistic looking colors which they then transferred to photographic Visible Human datasets. Lu et al. [38] used colored example images

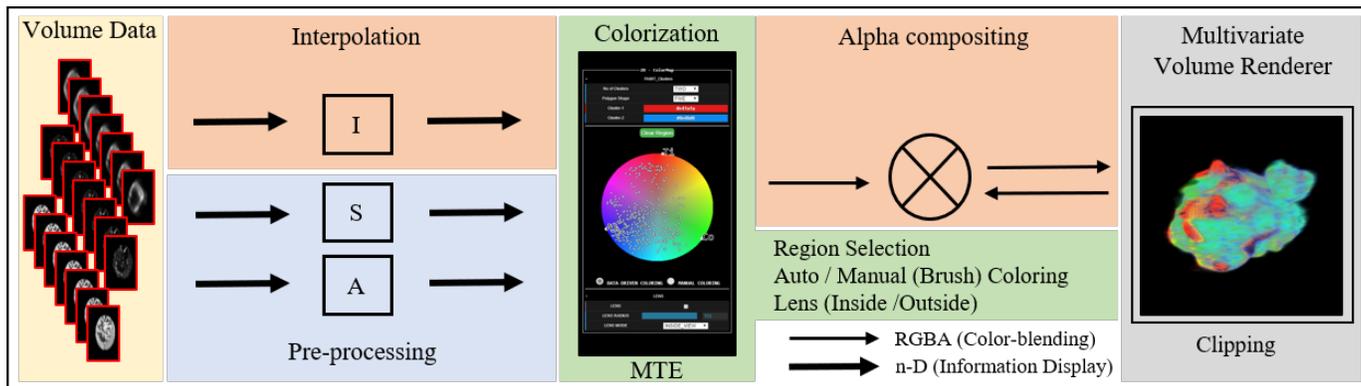


Fig. 2. System overview. A multivariate volumetric data set is loaded into the GPU pipeline. The pre-processing step sets up the information display. Process A computes the correlation among the channels to determine their spacing on the display’s boundary and then embeds the data into the circular information display. Next, S is a de-cluttering step. It down-samples the embedded points while preserving their spatial distribution. The rendering can now begin (process I) which streams the volume data through the information display, interpolates the colormap, and performs alpha compositing with the emerging volume display until all data have been streamed through. Users can brush or paint with the transfer function editor, change the parameters of the compositing pipeline, or perform spatial clipping to highlight volumetric structures of interest in the rendering.

to produce a 3D volume with a similar look using texture synthesis. Kindlmann et al. [39] used hue-balls for coloring and a 2D barycentric space of anisotropy for the opacity of data. It was later extended for use within barycentric colormaps by Kindlmann et al. [40]. Welsh et al. [41] use texture matching to transfer RGB color to grey-level volumetric data.

4 DESIGN GOALS

Upon embarking on this project we conducted a series of interviews with several domain experts in the area of nanoscale imaging, microscopy, and spectroscopy. The modalities they typically use are hard X-ray and transmission electron microscopy. We find that the goals of these scientists are quite general and will likely apply to other domains as well. As such our methodology is designed to be widely applicable to any multivariate volumetric data. Based on these interviews we derived five design goals:

DG1: Visualize all channels simultaneously. This was identified as a major requirement as it would address a main shortcoming of the current state of the art. It would allow scientists to better and more expeditiously appreciate the spatial composition and chemical heterogeneity in the data.

DG2: Enable semi-automated data-driven color labeling. Obtaining an automatically-generated meaningful first impression about the data overcomes cold start problems in data analytics. A data-driven color mapping scheme that could automatically highlight different chemicals and their compositions seems a suitable approach to achieve this goal.

DG3: Support selection and filtering to control display complexity. As discussed, aggregation of volumetric multivariate data into a colored 2D image is an inherently ill-posed task. A common way to deal with this problem is to provide a set of interactive functionalities that allow users to explore desired aspects of the data autonomously.

DG4: Enable manual labeling (brushing). Once users have gained a good understanding of the volumetric structures they require a method by which they can express these findings via concrete labeling. This could be achieved by a manual brushing tool operating directly on the data.

DG5: Make the tool platform-independent. Researchers use a variety of OS platforms and file systems. To achieve wide applicability our tool should be browser-based.

5 APPROACH AND IMPLEMENTATION

Figure 2 gives an overview of our overall system. The first stage is the formation of the circular RadViz-inspired information display which consists of two steps: (1) place nodes representing the channels in proper positions on the disk’s boundary, and (2) embed the multi-D channel data into the disk’s interior (Figure 2 (A, S)). At this point, we have a multivariate information visualization of the volume data (DG1). The HSL colormap fused with the circular information display establishes the link to the volume rendering pipeline by ways of its function as a multivariate transfer function editor (MTE). We can now engage in the second (iterative) stage – the exploratory volume visualization. This stage uses the information display mapping to assign colors and opacities to the samples interpolated from the volume (Figure 2 (I)), supported by various interactive operators.

5.1 Background: RadViz Inspired Information Display

In the following, we provide a brief summary of the relevant features of the colormap-fused information display developed by Cheng et al. [42] which serves as the backbone of our multivariate volume visualization pipeline. Pictured in Figure 1(B) and in the box labeled ‘MTE’ in Figure 2, it consists of a circular multivariate information display fused with a radial-slice color map, such as HSV, HSL, or HCL. Cheng et al. used this interface to interactively colorize multi-channel 2D images, acquired either with multi-spectral cameras or generated from multivariate choropleth maps. There was no support for volumetric data which greatly simplified the mitigation of the aforementioned aggregation problem (see Section 1).

By default, the channel nodes on the display’s boundary are equidistantly spaced such that each node will map to a distinct “primary” color. Alternatively, the nodes can be spaced in terms of the correlation similarity of the channels. Choosing this spacing results in a configuration where

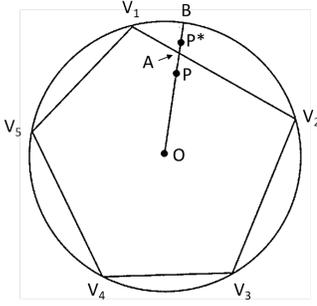


Fig. 3. The polygonal mapping region of RadViz and the extension to a circle to enable the full use of the HSL color space (redrawn from [43]).

correlated channels are mapped to similar colors. Here, one would first calculate a correlation matrix of the channels and then compute an efficient ordering of the nodes via a Hamilton Cycle algorithm [43]. To calculate the optimal spacing between the nodes on the MTE circular boundary we use the metric $(1 - \rho_{ij})$ where ρ_{ij} is the correlation of attribute i and j as follows:

$$s_{ij} = \frac{1 - \rho_{ij}}{\sum_{k,l \in HC} (1 - \rho_{kl})} s_{MTE}$$

Here, s_{MTE} is the MTE's circumference and s_{ij} is the distance between two attributes i, j adjacent in the Hamilton Cycle HC . The spacing so obtained groups similar attributes close together, which are then assigned similar colors. This is in some sense a dimension reduction, saving any distinct primary colors for more independent attributes.

With each of the channels assigned a specific position on the circle's boundary, RadViz uses the following equation to map a data point to the circle's interior:

$$P_x = \sum_{j=1}^n a_j v_j \text{ where } a_j = w_j / \sum_{k=1}^n w_k$$

P_x is the location of the mapped data point, the w_j are the values of the n -D data vector, and the v_j are the spatial locations of the channel nodes on the circle boundary. This mapping embeds the points into a polygonal shape defined by the locations of the channel nodes. To accommodate the full circular disk space of the color map, Cheng et al. [42] devised a stretching formula where a point P inside the polygon is translated to a point P^* defined by $OP^* = OP \cdot OB/OA$, where O is the center of the circle and A and B are the intersection points of the line OP with the polygon and the circle, respectively. This is illustrated in Figure 3.

We use the above equation for two purposes: (1) to map the original data points for the visualization of their multivariate distributions, and (2) to lookup/interpolate the RGB colors for the vectors generated by the sampling process in the multivariate volume rendering.

5.1.1 Dealing with Ambiguities

The first potential ambiguity stems from the circumstance that both RadViz mapping and interpolation only observe the *relative* proportion of a data point's channel, as opposed to its absolute value (the magnitude, Euclidean norm). As a remedy, we provide a set of sliders (bottom two sliders in (Figure 1(E))) in the volume rendering interface that allows

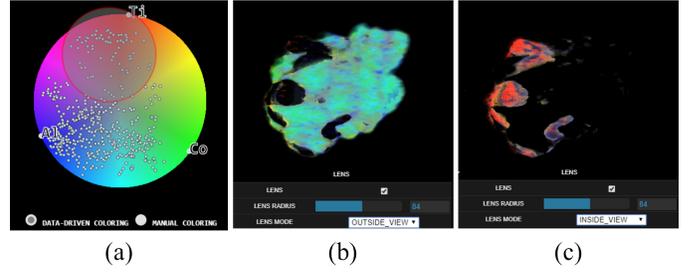


Fig. 4. Lens filtering using the *LiCoO2* dataset described in Section 6.3. (a) the information display with the lens, (b) the volume rendering of the samples covered by the lens and (c) of the samples outside of the lens.

users to choose a magnitude range of interest which renders all samples that fall outside this interval invisible. This is similar to opacity-culling in standard volume rendering.

The second ambiguity arises because the RadViz mapping is not unique; vectors with different value compositions may map into similar neighborhoods. Cheng et al. [43] devised three jointly applied techniques by which this ambiguity can be markedly reduced. Two of these require a non-linear optimization of the layout; they are not feasible in our case since they do not allow for the color lookup which requires a linear mapping. The third remedy is the aforementioned correlation-based channel node positioning on the RadViz boundary. It turns out that this can already resolve a significant amount of ambiguity since it maps similar colors to similar channels and consequently allocates more space to other channels which reduce the potential of dissimilar points being mapped into the same map areas.

5.2 Our RadVolViz Color Labeling Scheme

The need to support volumetric data required several non-trivial extensions to Cheng et al's original colormap widget. It resulted in a new colormap widget we call *Multivariate Transfer Function Editor (MTE)* which has two main new features: (1) several interactive operators that allow users to control the visual properties by which the channel mixtures are visualized in the volumetric display (DG3), (2) a manual brushing/coloring facility by which users can precisely color-label the structures identified during the prior exploration (DG4). The two color labeling modes (data-driven and manual) can be selected via the radio buttons below the MTE (see Figure 1(B)).

DG1 and DG2 were already largely supported by the original colormap widget; they mainly required extensions of the volume rendering engine we utilized (see below), a sampling scheme to prevent the overplotting that resulted from extending the data from images to volumes, and the selection of an appropriate color space for the color map. We now describe each of these components in closer detail.

Overdraw Management: The number of data points in a volumetric dataset can easily overwhelm the MTE information display and cause a high amount of overdraw and clutter. To manage this overdraw, we follow the information display mapping step by a stratified downsampling operation. The sampling is done such that the mapped points preserve the spatial distribution of the original volume data. For this purpose, we subdivide the display's disk into an array of square-shaped bins. The bin size can be set via the *BIN* widget shown in Figure 1(G) and the overall

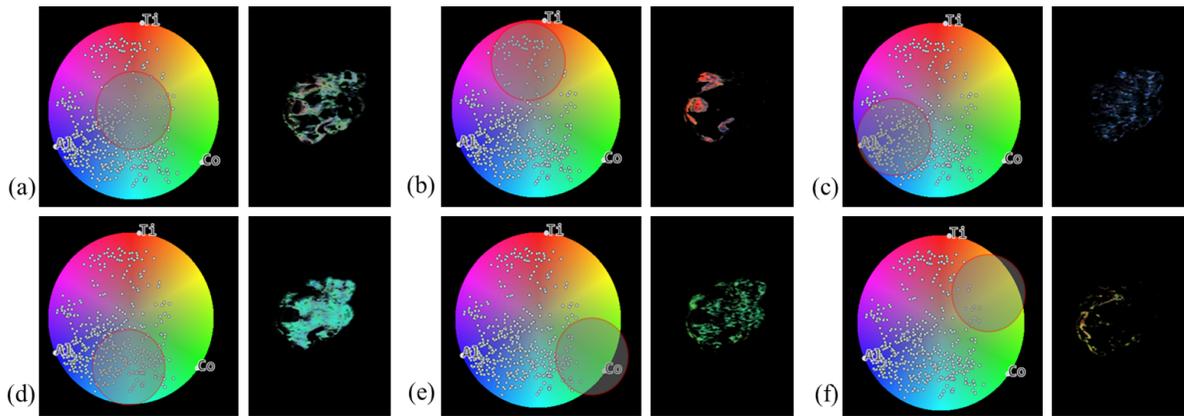


Fig. 5. Rendering results due to various lens placements using the *LiCoO2* dataset described in Section 6.3: (a) relatively homogeneous samples, (b-f) samples with relatively high Titanium (b), Aluminum (c), mixtures of Aluminum and Cobalt (d), Cobalt (e), and mixtures of Titanium and Cobalt.

number of data points in the colorspace m can be set using the *POINTS* setting. The stratified sampling in each bin is determined from the total number of data points, M . Say, the number of desired points is m , then the stratification rate $r = m/M \cdot 100\%$. Each bin is randomly sampled such that $r\%$ of the original points remains. Bins with just a few points are left unsampled to maintain awareness of outliers. In practice, we use a grid of 6×6 bins which in conjunction with the aforementioned sampling scheme typically results in about 500 points being left on the map.

Data-Driven Color Labeling: The MTE information display (1) informs the user about the multivariate distributions of the volume data, and (2) provides a color for each multivariate (ray) sample generated by the volume renderer (DG2) using the interpolation scheme outlined in Section 5.1. To determine the color we chose to utilize a color map derived from the HSL color space. While Cheng et al. used the HCL color space, our switch to HSL was mainly driven by our new application to volume data: (1) HSL has a significantly more pronounced white point in the center where voxels with equal values in all channels are mapped into (see Figure 1(B)); we found that this helped to better differentiate these special voxels in the volume display. (2) HSL does not have color regions outside the circle; the HCL space has such regions since it is a cut-out of the LUV space. While Cheng et al. allow users to stop and expand the HCL color space to the outside LUV regions for greater color contrast, we experimentally found that this can lead to inconsistent displays in the volume rendering.

The HSL space has a double cone topology and turns into a disk if the lightness L is fixed. Hence, the MTE will result in the ray sample’s Hue H and Saturation S while we are still free to determine L . We have experimented with different settings and found that typically $L=0.55$ showed the best contrast. All images in this paper use this level.

5.2.1 Interactions and User Interface

Although our tool is data-driven and configured automatically (DG2), it includes a variety of interactive features to allow scientists to explore the data (DG3) and then report the results (DG4). We describe these utilities next.

Lens Filter: In order to control the mapping of data points for any particular channel or multiple channels together, we devised a variety of filters that are blended into

the information display. One of these filters acts as a lens. The lens is controlled by three parameters, namely *LENS*, *LENS RADIUS*, and *LENS MODE* as shown in Figure 1(D). *LENS* is used for selecting/deselecting the lens operation in the information display, and the *LENS RADIUS* specifies the radial extent of the circular lens region in the colormap. The effect of the lens-selected region is determined by *LENS MODE* i.e. *INSIDE_VIEW* and *OUTSIDE_VIEW*. While the former is used to emphasize the effect of the region selected in the volumetric structure (voxels selected inside the lens), the latter isolates it.

Figure 4(a) gives an example of a lens selecting a subset of points mapped in the MTE. In Figure 4(b) the points within the lens are volume-rendered using the lens *INSIDE_VIEW* mode, while in Figure 4(c) the points that lie outside of the lens are rendered using the lens *OUTSIDE_VIEW* mode. Essentially the two modes produce two complementary renderings of the volume which can be convenient in volume exploration tasks. The lens can be moved around the colormap via simple mouse drag interactions which allow for easy material or material compound focused volume explorations. Some examples are shown in Figure 5 and further explained in the caption.

Correlation-Driven Mapping: Our tool also enables domain experts to dig deeper into the information display and explore correlations both among the sample points and among the channels and see these emphasized in the visualized volume. The *MODE* widget as shown in Figure 1(G) has three options, *RGB*, *GBC* and *DGBC*. *RGB* is an option reserved for three-channel density data; each channel is mapped to R, G, or B; a trivial extension of the 1D transfer editor. Option *GBC* is the mapping achieved when the variables are spaced uniformly around the information display boundary, while option *DGBC* places them spaced apart in terms of their correlation (see Section 5.1).

Figure 6(a) shows the rendering outcome when channels *Ti*, *Co*, and *Al* are directly assigned to the R, G, and B channel, respectively (essentially three separate 1D transfer functions), whereas Figure 6(b) is the rendering when *Ti*, *Co*, and *Al* are assigned to an equispaced three-channel MTE (similar to Figure 4 but without a lens). We confirm that Figure 6(a) and 6(b) show the same structures at similar colors, but Figure 6(b) seems to have better contrast, both

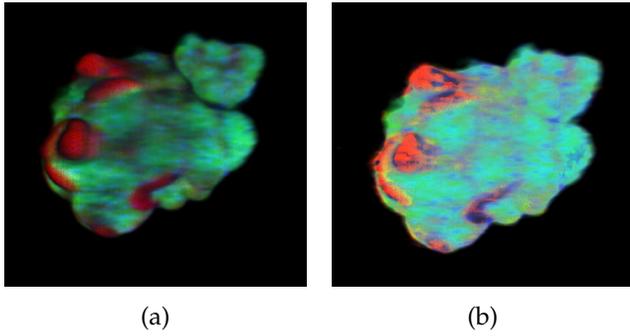


Fig. 6. Visualization of the 3-channel $LiCoO_2$ dataset described in Section 6.3. The three channels are Ti, Co, and Al. In (a) each channel is directly assigned to a Red, Green, or Blue color, while in (b) the MTE-based color mapping is used (in equispaced GBC mode). We observe that the MTE-based mapping shows significantly more contrast.

in the interior and the exterior. This might be due to the more spread out data distributions in the circular colormap.

Finally, we compare Figure 6(b) which uses a GBC-based colorization with Figure 7 which uses DGBC. We observe that DGBC gives *Ti* and *Al* more similar colors due to their close correlation but spaces *Co* far apart since it is not correlated to either of them at all. As a result, *Co* stands out significantly more in the corresponding volume visualization and the *Al* speckles in it are much more apparent.

Manual Color Labeling: This mode gives users the ability to refine the color labels of material compound clusters identified in the data-driven mode (DG4). The manual brushing interface allows users to coordinate the labeling with the real-time spatial feedback provided by the volume display. We noticed that the scientists often started a session with the data-driven interface (Figure 8(a)) to get an overall idea about the dataset and then switched to the manual brushing interface (Figure 8(b)) to elucidate the structures precisely. An example for the results of such an activity is shown in Figures 14(a) and (c).

The manual color labeling functionality allows users to assign any desired color to a selected cluster of voxels. Two widgets assist in this activity, *No of Clusters* and *Polygon Shape*. The former lets users select the number of clusters to brush, while the latter assists them in marking the boundaries of the brushed clusters (Figure 1(A)). Marking the boundaries in the desired polygonal shape is accomplished by simple mouse clicks to cover the area of interest (clusters). Figure 8(b) shows the widgets used to select a cluster's

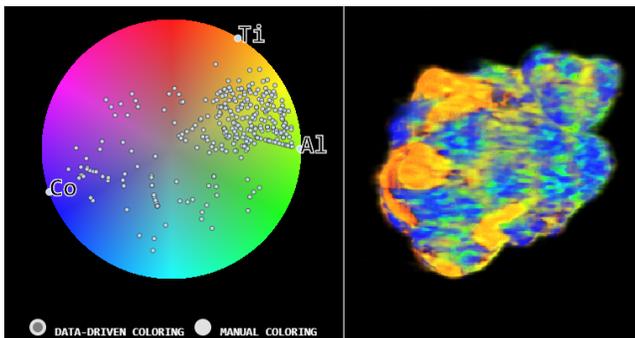


Fig. 7. Visualization of the 3-channel $LiCoO_2$ dataset. (Left) the MTE in the DGBC (correlation-based) arrangement of the channels on the MTE boundary, (right) the corresponding volume rendering. The DGBC rendering better differentiates Co and so provides higher contrast for Al.

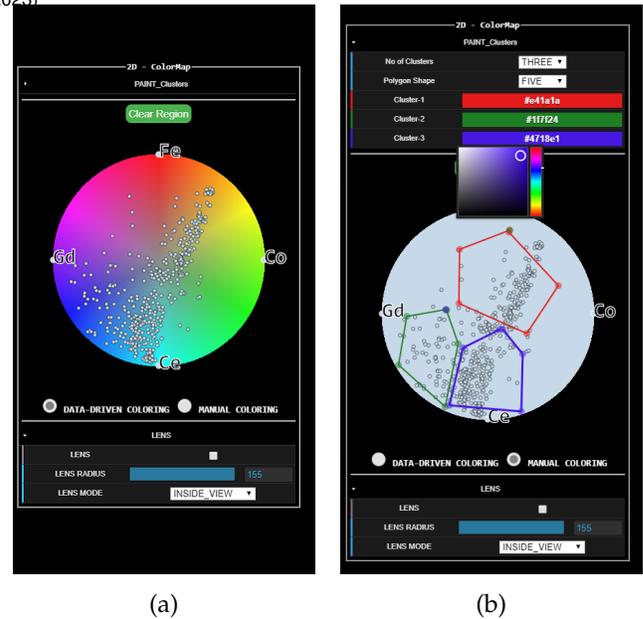


Fig. 8. The two colorization modes, both showing the RadViz (GBC)-type embedding of a sampled subset of the volume data: (a) Automated mode with HS colormap; voxels mapped to a specific location will be given the color associated with that location. (b) Manual (brushing) mode with no colormap; the user has selected three regions and voxels mapped into these regions will be tagged in the associated color.

information and the polygonal shape used to define it. To label a cluster, users can pick from a set of default colors or choose a color from a pop-up color selection pad.

Color Map Rotation: Users might have a preference for the colors assigned to certain channels. For this purpose, we provide a *ROTATE* widget option as shown in Figure 1(G) by which users can rotate the channels around the MTE boundary and so change the mapping. The scientists we worked with very much liked this flexibility as it enabled them to enforce common conventions in their field. Figure 14(a) is the result of *Ce* being projected to the bottom of the circular colormap, whereas Figure 14(b) is an outcome of *Ce* being mapped to the right of the colormap. This paradigm can also benefit individuals blind to certain colors. They can simply rotate the map to avoid these colors or color combinations when generating the volumetric image.

5.3 Our Multivariate Volume Renderer

We chose a post-classified raycasting paradigm where a color is assigned to a ray sample after interpolating the multivariate volume. This rendering pipeline is fairly straightforward to implement on the GPU. Once the volume file is loaded into GPU memory as a texture, it is sampled along the rays at regular intervals. Volumetric information along with other parameters provided by the MTE is collectively used to generate a ray sample's color and opacity which is then composited in the fragment shader with the respective values aggregated so far. We use front to back compositing which allows the cumulative opacity to be tracked and the computation along a ray terminated when the accumulated opacity gets close to 1.0.

5.3.1 Interactions and User Interface

We added several interactions to the renderer to help users explore the data; many of these are tied to MTE-based interactions. We describe these now in the following.

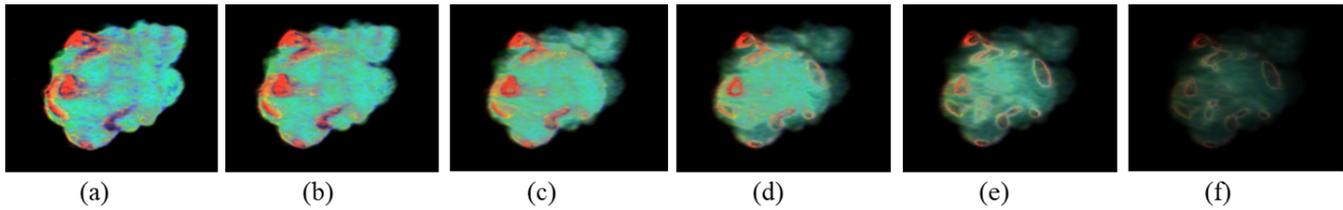


Fig. 9. Different opacity settings for the $LiCoO_2$ dataset: (a) 0.99, (b) 0.51, (c) 0.25, (d) 0.15, (e) 0.1, (f) 0.04. Lower opacity settings allow the visualization of interior structures. One can simply move the opacity slider back and forth to appreciate the spatial relations of these structures.

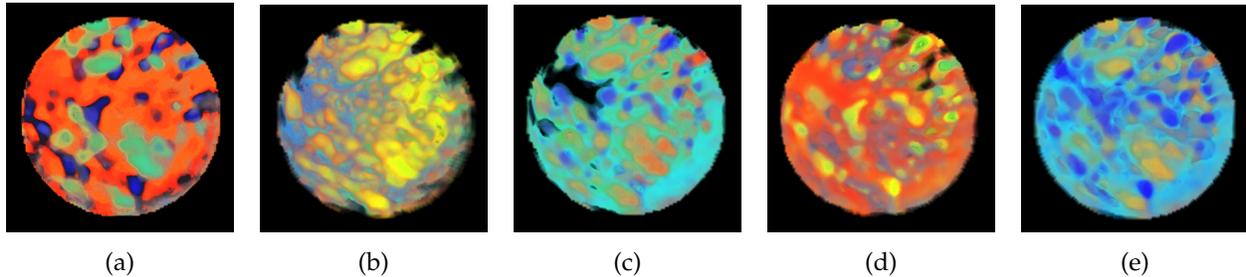


Fig. 10. Renderings of the Hard X-ray dataset achieved with data-driven alpha-blending. In all images the colors are generated with the MTE color mapping as described (a) using Ce as the alpha channel, (b) using Co, (c) using Fe, (d) using Gd and (e) using the average of all attribute channels.

Data Driven Alpha Blending: Opacity weighting is a powerful technique for emphasizing volumetric structures of importance. We provide a widget *OPACITY* as shown in Figure 1(E) where users can set the alpha-value used in the compositing at all steps of the ray. Figure 9 shows a set of example renderings achieved with different opacity settings. We observe that setting the opacity to lower values allows an efficient visualization of interior structures.

Our application also allows users to choose one or more of the volume channel attributes as the opacity channel. Figure 10 shows an example using the multi-modal X-ray data set described in Section 6.2. It shows renderings produced with either one or all volume attributes being used for the alpha compositing. Figure 10(a) is the output obtained when all of the four material channel data are mapped to color and *Ce* is used for the alpha computation. Equivalent outcomes for other attributes serving as opacity are shown for *Co*, *Fe* and *Gd* in Figure 10(b) - (d) respectively. In each case, the generated image shows a resemblance to the texture structure of the respective channel used for alpha weighting. We found that this mode can be useful to recognize the spatial organization of specific channels.

To consider the effect of all of the attributes on the displayed volumetric structure, we perform alpha blending of the average of all attribute channels, using an average sum of these for alpha. Figure 10(e) shows an example of this using the same dataset. We can see that the patterns present in Figure 10(a) and (c) are missing in Figure 10(b) and (d), but are present in Figure 10(e). We use the combined channel alpha-average as the rendering mode throughout the paper.

Clipping: To allow users to explore the inner composition within cross-sections of the volume, we added clipping as an additional functionality. Clipping can occur along both an axis and a general plane. As shown in Figure 1(H), a drop-down widget lets the user select either the *Axis* or the *Planar* option for clipping. Figure 13(c, d) shows an example of planar clipping where the clipping is perpendicular to the plane formed by the values assigned to *X CLIP*, *Y CLIP*, and *Z CLIP*.

5.4 System Implementation and Scalability

We implemented our system as a web-based framework to gain platform independence (DG5). We used WebGL and a higher-level API built on WebGL, Three.js. The major advantage of WebGL is that it runs under the control of the browser and has no direct communication with the underlying operating system. We store the multivariate volumetric data as 4-channel RGBA image textures. If the number of channels in the visualized compound of materials is greater than four we generate another image texture. The texture image loaded with the channel information is fed into the WebGL framework for multiresolution interpolation, then mapped to the HSL colormap for interpolation, and the resulting color is composited with the evolving image. All of these operations are solely GPU-based; not much complexity is involved in the rendering pipeline.

Since the user interface design capabilities of WebGL are limited we implemented the MTE in javascript aided by the D3.js library. Hence, all user interactions with the MTE, such as brushing and lensing occur on the client-side CPU. On the other hand, the volume-side interactions, such as clipping, occur on the GPU. To inform the GPU rendering of MTE interactions we keep a shadow MTE process on the GPU. Any CPU-side user interaction with the MTE is rapidly communicated to the GPU shadow process and is reflected in the next rendering pass. There is only negligible overhead associated with this tandem process and so the rendering framerate is not visibly affected.

Scalability: The datasets we studied with our domain collaborators typically had 3-4 channels. To test scalability to a greater number of channels we tested our system with a 9-channel dataset we obtained from another collaborator. The dataset was acquired with a Micro X-ray fluorescence system and captures the elemental distributions in spheroids of human neuroblastoma cells. These cells are composed of nine elements: *Zn*, *Si*, *Fe*, *Ni*, *Cl*, *Ti*, *Ca*, *Cr*, and *Cu*. The volume visualizations we created can aid in studies investigating the role of trace elements in Neuroblastoma [44]. The results are shown in Figure 11. The number of frames per

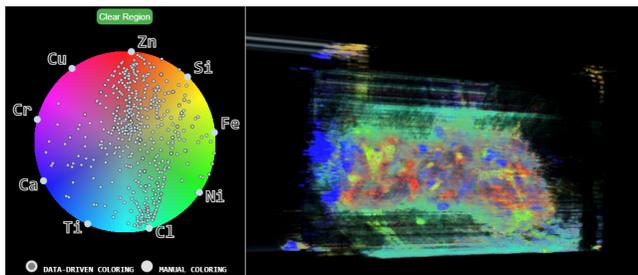


Fig. 11. Scalability of our system tested with a neuroblastoma cell dataset with 9 elements: Zn, Si, Fe, Ni, Cl, Ti, Ca, Cr, Cu. (Left) the MTE configured into 9 channels, (right) the associated volume rendering. One can clearly observe distinct substructures, distributions, and associations. The MTE can be used to look up their chemical compositions.

second was in the range of 45 - 60 in all of the three case studies we undertook. We did not notice a significant drop in framerate.

The theoretical limit of the MTE with regards to the number of attributes it can represent has already been assessed by Cheng et al. [42] using perceptual arguments for the CIE LUV color space. There, assuming the just noticeable difference (JND) of about 13 the number of distinguishable primaries in a ring layout with uniform spacing is about 35. While we use a less perceptual color space even half that number is not overly likely to occur in our current target field, even when additional channels are derived from the present data such as gradients or time. This number is potentially even higher with the correlation-based DGBC channel node spacing on the MTE boundary.

6 RESULTS FROM FOUR STUDY APPLICATIONS

Throughout the course of system development, we had several meetings with two teams of material scientists. Each scientist had ten or more years of experience in the field of material science. The meetings were separate since the two teams work at different geographic locations, with different imaging modalities, and have different research goals and approaches. Yet, both teams share a common goal, namely, finding better and more cost-efficient materials for batteries.

Each meeting session was about 1-2 hours long and was conducted over the web. Following each session, we would then expand our tool with new functionalities that we felt addressed the issues raised. Each new version we produced usually not only satisfied these issues but also stimulated new ideas, and occasionally the outcomes of these sessions were surprising and unanticipated. Eventually, the tool met all, or at least the great majority, of the domain scientists' expectations. The results and insights presented in the following are the results of the final evaluations of our tool with three of the most frequently used datasets.

6.1 Chemically Sensitive Electron Tomography Study

The subject of this study was a four-channel lithium-ion battery volumetric dataset, acquired via chemically sensitive electron tomography. Of interest was the cathode material, composed of Lithium (*Li*), Manganese (*Mn*), Nickel (*Ni*), Cobalt (*Co*), and Oxygen (*O*). We used our tool to diagnose the cathode material's synthesis process and the distribution of materials in the imaged sample. The exploration was

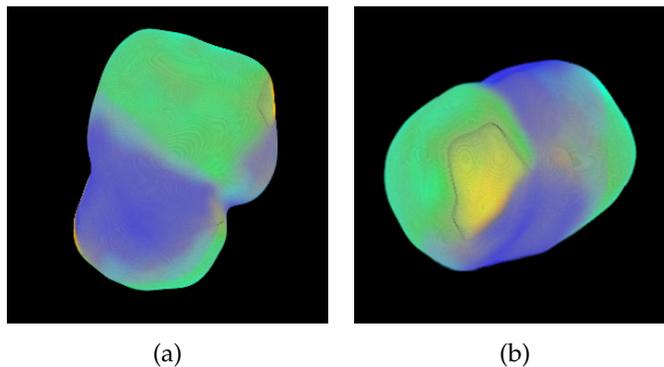


Fig. 12. Multivariate volume rendering of the four-channel lithium-ion battery dataset. The opacity was set to reveal the sample's surface. View (a) shows that the sample is rich in *Mn* (green) and *NiO* (blue). View (b) shows that there is a fairly high presence of desirable *Co* and *Mn* (yellowish) but also the less desirable *NiO* (for the MTE's color assignment refer to Figure 2(B)).

motivated by the quest to find a substitute for Cobalt; a material that could deliver more charge or energy and at the same time would be economically more viable.

Some of the composites we found were *Ni*, *Mn* and Cobalt Oxide. As mentioned, the overarching goal is to replace the large portion of *Co* with a mix of *Mn* and *Co*; it would improve the charge density which is the measure of how much charge one can pull out per unit weight. Most electric cars, such as Tesla, use a combination of materials in their batteries, prominently *Li*, *Ni*, *Mn*, *Co*, and *O* because they are cheaper and have higher charge density. Our tool's role was to help engineers determine whether the cathode materials were synthesized properly or not, and if not, which of the synthesized elements was not desirable.

The data for this case study was composed of an image stack of four elements: *Mn*, *Ni*, *Co*, and *O*. It lacked *Li* since the engineers were unable to map out the *Li* in the compound. The images were of size 204×204 ; there were 128 slices in total and the voxel size was $1.5 \text{ nm} \times 1.5 \text{ nm} \times 1.5 \text{ nm}$. The dataset fully reconstructed the cation and anion distributions of technologically relevant primary particles with a material density of 4.93 g/cm^3 .

Figure 1 (C) shows one of the visualizations we created together with the scientists. All visualizations we demonstrate in this section use the colormap and information display shown in this figure. The various visualizations that were produced revealed the spatial distribution of all materials in the volume. In general, for a synthesized material to be desirable and suitable for use in lithium-ion batteries, there should be chemical homogeneity. As shown in Figure 12(a), one side of the surface is mostly greenish in color which shows an abundance of *Mn*, whereas the other side of the volume as shown in Figure 12(b) has a yellowish distribution which suggests the presence of desirable *Co* and *Mn*-rich regions in the material. But it also has a (smaller) blueish region, which is mostly Nickel Oxide (*NiO*). This distribution is also seen in the MTE display of Figure 1(B).

The volume visualization helped the scientists to quickly recognize that this cathode material's surface was chemically heterogeneous which they say is not overly desirable. The early diagnosis was made possible through the simultaneous visualization of all four channels; it can accommodate all chemical compositions, which is the primary goal of this

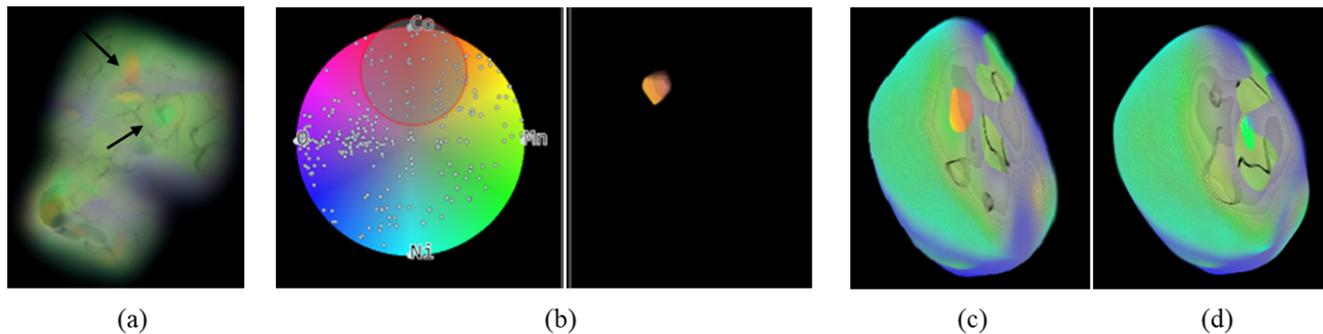


Fig. 13. Multivariate volume rendering of the lithium-ion battery sample. (a) Setting a low opacity threshold (0.05) reveals interesting interior structures (indicated by arrows). (b) Using the lens confirms that one of the embedded structures is metallic cobalt. (d-d) Subsequent clipping then fully exposes their presence, namely (c) metallic cobalt and (d) nickel-rich magnesium in the core of the sample. Both discoveries point to deficiencies in the sample's synthesis process. (Not all images were taken from the same viewpoint.)

analysis. The images readily show that there is no strong presence of *Ni*, while the presence of *Mn* and *CoMn* is quite rich, which is the desired structure for this cathode material.

Next, the scientists wanted to explore the inner composition of the volume. They first used the opacity slider to reveal the inner structures and visually detected an orange-yellow cluster as well as a greenish cluster, both faintly shining through (see arrows in Figure 13(a)). They moved the lens to critical regions on the MTE to confirm the chemical substance of these structures. This is shown in Figure 13(b) for the orange-yellow cluster which is identified as metallic cobalt. They then used our tool's clipping feature to "excavate" these findings. On that journey they were able to fully identify both the orange-yellow *Co*-rich cluster embedded deep in the material sample (Figure 13(c)) and the greenish deeply embedded *NiMn* cluster (Figure 13(d)). This was a very insightful finding since it is undesirable and not expected to see any pure metallic clusters in a sample. The scientists concluded that there was a limitation in oxygen diffusion in the material's synthesis process. It meant that both metallic cobalt and nickel-rich magnesium existed in the cathode oxide materials even though the materials were calcined in oxygen. It seemed highly certain that there was an oxygen deficiency in the process of calcination, a high temperature heating process used for thermal treatment in the absence or limited supply of air or oxygen and applied to ores and other solid materials to bring about a thermal decomposition. The visual examination of our renderings helped the scientists diagnose that the synthesis process had problems since the generated materials were not synthesized as intended.

6.2 Phase Detection in Multimodal Hard X-Ray Data

This use case stems from our most recent session with the other team of material scientists we have been working with. They seek to investigate the effectiveness of multimodal imaging at the nanometer scale and the formation of grain boundaries formed by complex interfacial characteristics. Analyzing these images helps them observe phenomena related to energy conversion and storage in solid oxide fuel cells [45].

In this use case, the primary goal has been to learn about the elemental diffusion and phase separation which can lead to the formation of additional material phases at the grain boundary. The formation of additional material phases

is believed to play an important role in determining the ultimate electron and ion transport properties of the imaged material [46]. The data consisted of fluorescence slice images of Cerium (*Ce*), Cobalt (*Co*), Iron (*Fe*), and Gadolinium (*Gd*); the slice image size was 150×150 pixels and there were 90 slices in total. Platinum (*Pt*) serves as a foreign element for sample preparation but is not considered in this analysis.

Our collaborators told us that with the help of their current tool [46], they were able to detect three different phases by analyzing the fact that *Gd* and *Ce* are not completely correlated, and neither are *Co* and *Fe* which led to the formation of two material phases throughout. Having no correlation among these two phases led to the formation of a third phase which they reported in their analysis. However, finding these material phases was tedious with current tools, especially when the number of elements exceeded three and they could no longer resort to RGB encoding.

With the help of our tool, they were able to find new material phases quite easily. This is evident from Figure 14(a) which is dominated by two material phases, visually encoded in Cyan and Blue, respectively. We used the equidistant GBC layout in the MTE to achieve a distinct coloring for the four elements in the volume display. The visualization also readily reveals the presence of the third phase, colored orange and distributed throughout the volume.

The lack of correlation between these materials leading to the generation of the current phases is also evident in the visualization generated with the correlation-based (DGBC) layout. Shown in Figure 14(b) we quickly see that *Co* and *Fe* as well as *Gd* and *Ce* are well separated on the MTE boundary. This is consistent with previous analyses [46].

Apart from showing consistency with previous findings, our tool also gave an interesting new insight, namely the presence of a fourth phase, colored in a faint magenta in Figure 14(a), indicated by a circle. However, this phase is not clearly visible and needs an extra careful observation to become more evident. It is because the material responsible for this phase is present at a smaller ratio which leads the corresponding voxels to map somewhere near the center of the HSL colormap. As a consequence, the points are desaturated and become achromatic (Gray).

To address this issue, we switched to the manual coloring mode. It allows an assignment of similar color values to all selected voxels, making it easier to highlight the voxels nearer to the center. Highlighting such voxels in this case helps in investigating the formation of phases composed

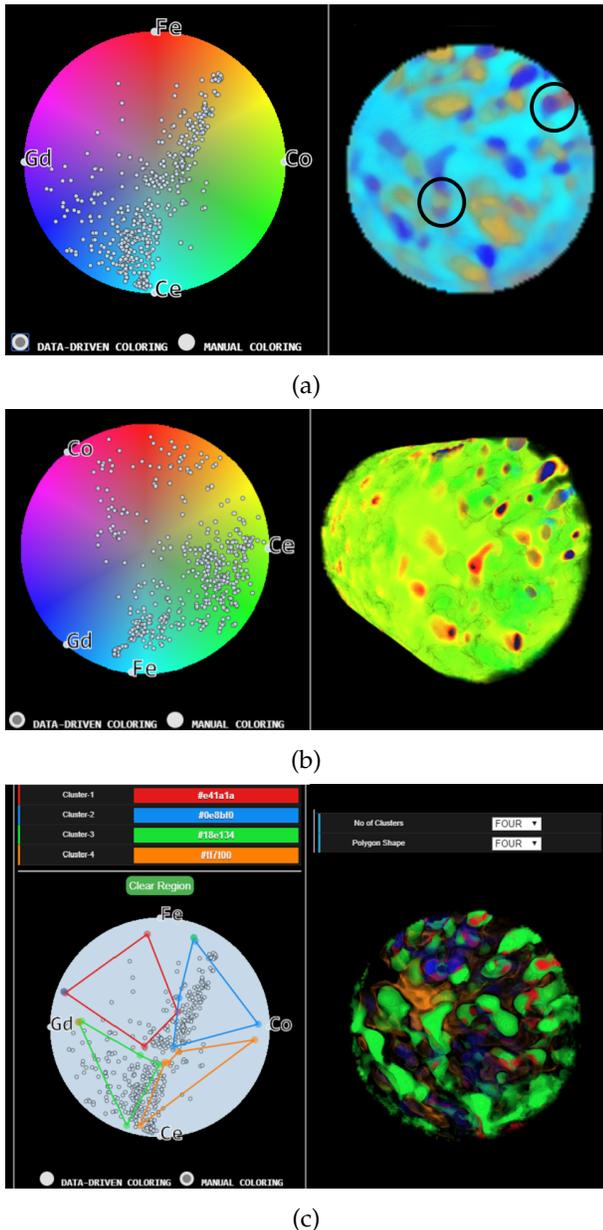


Fig. 14. Hard X-ray dataset with four materials Ce , Co , Fe and Gd . (a) Visualization using the GBC layout shows the presence of 3 phases and even indicates the presence of a 4th phase in a faint magenta (circled). (b) The DGBC layout explains the reason for the existence of phases, i.e. the lack of correlation between materials. (c) Visualization after manual coloring highlights the four phases discovered in (a) more prominently.

of lesser dominant materials. The outcome of this manual coloring can be seen in Figure 14(c) where the fourth phase is now clearly evident in red color.

6.3 Trace Doping of $LiCoO_2$ in a Lithium-ion Battery

We also evaluated our tool on a dataset of $LiCoO_2$ which is a dominant cathode material for lithium-ion batteries. The goal of this study was to investigate the elemental distribution of Ti - Mg - Al co-doped $LiCoO_2$ (TMA-LCO) to learn more about its surface stability which is important for building cathode materials with higher voltage charging capacity and larger reversible capacity. At higher voltage, internal strain builds up, and oxygen loss brings irreversible phase transitions. These lead to surface instabilities which in turn lead to serious performance degradation of $LiCoO_2$.

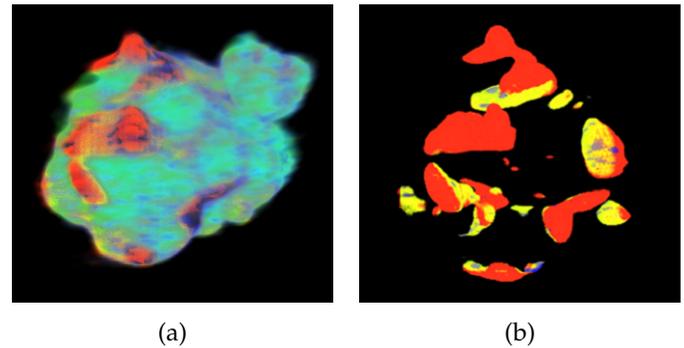


Fig. 15. $LiCoO_2$ dataset (a) Volumetric structure for Ti (red), Co (turquoise), and Al (dark blue) highlighting the material composition and locations in the volume. (b) The Ti component is highlighted in red using the lens and the magnitude min-max filtering operations to show its separate presence on the particle surface.

We visualized the foreign element doping strategy which has been demonstrated to be highly promising and effective for the improvement of the electrochemical performance of $LiCoO_2$ [4]. To study the spatial elemental distribution, the scientists did a fluorescence mapping of the doped material and collected stacks of Al , Co , and Ti (Titanium) images for analysis. The images were of size 100×100 and there were 100 slices in total. The volume visualization generated from this dataset in Figure 15(a) shows the elemental distribution of the three elements using GBC mapping (the MTE is shown in Figure 4). It is evident from Figure 15(a) that Al (dark blue) and Co (turquoise) are distributed throughout the entire particle which is consistent with the findings from experiments by Zhang et al. [4].

An interesting finding is the distribution of Ti (red) which appears largely separated into smaller groups on the surface, as is clearly shown in Figure 15(b). This figure is a refinement of Figure 4(c), created by first placing a lens over the area around the MTE's Ti channel node and then restricting the volume rendering to high-magnitude samples by the magnitude min-max slider (Figure 1(E)). The separated presence of the Ti -rich phase on the surface ascribes to the long-term stability of $LiCoO_2$ by effectively reducing the lattice breathing and making it robust against lattice strain and particle fracture. The conclusion is consistent with experiments done by Zhang et al. over many TMA-LCO particles, confirming the heterogeneous distribution of Ti .

6.4 Hurricane Isabel Dataset

To show the generality of our tool, we also visualize the Hurricane Isabel dataset from a simulation in 2003. It includes data for a $2,139 \text{ km}$ (east-west) \times $2,004 \text{ km}$ (north-south) \times 19.8 km (vertical) volume. The simulation data is from the National Center for Atmospheric Research (<https://ncar.ucar.edu/>), and it was used in the IEEE Visualization contest 2004 (<http://vis.computer.org/vis2004contest>). We used 4 dimensions, i.e. pressure, temperature, wind speed magnitude, and water vapor mixing ratio (QVAPOR). The other variables had only low variability. The panels and caption in Figure 16 illustrate how valuable insight can quickly be obtained with just a few lens and slider interactions.

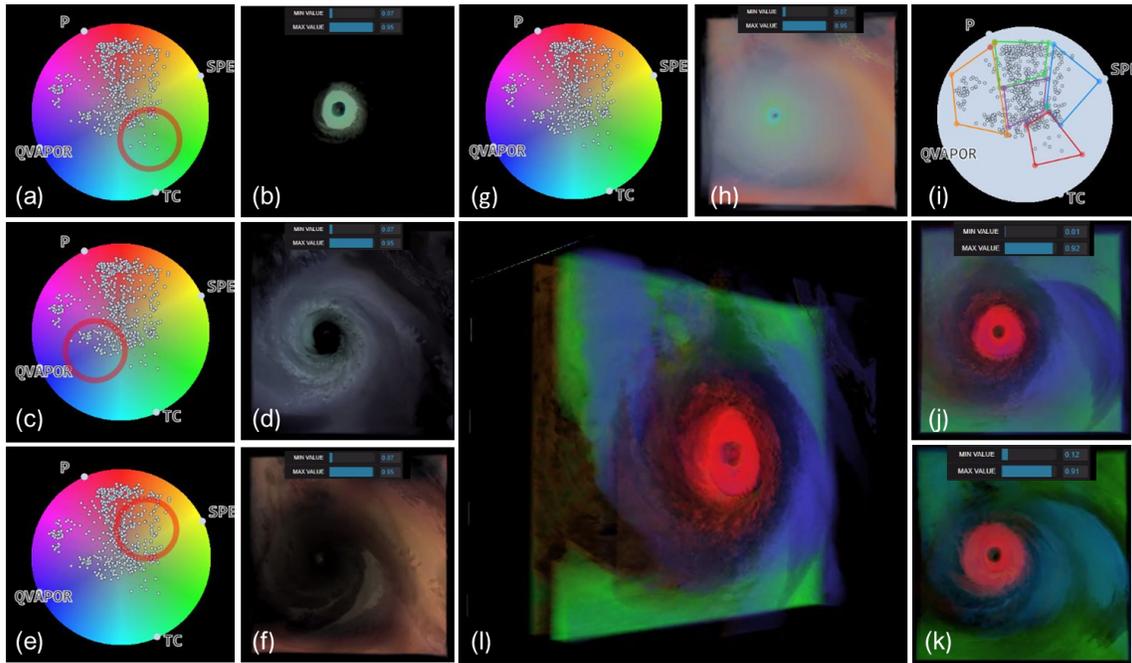


Fig. 16. Hurricane Isabel dataset. During setup, we noticed that of the 13 variables, only 4 (temperature, speed, pressure, and vapor) had significant variation and so we focused on these 4 to make the best use of the MTE’s real estate. (g-h) MTE and generic rendering with the min/max value ranges indicated by the two slider settings on top of the image. Even from this “out of the box” rendering it is already apparent that the eye is dominated by temperature (light blue to green color) while the eye’s periphery adds yellow to the green indicating a growing dominance of speed. Indeed, the wall around the eye is typically the storm’s strongest part and it pulls in warmer ocean water that increases the temperature in the eye. As the water is pulled up, moving further away from the eye the vapor is increasing (blue-tone colors) and finally the pressure (red-tone color). (a-f) Lens-driven renderings where we more clearly observe the high temperature in the eye (a-b), a classic cloud rendering of the vapor in grey-blue (c-d), and the pressure-dominated outer periphery in red-orange (e-f). (j-l) Colorizations using the manual brushings shown in (i) with different min-value slider settings; the regions with speed-dominated data are colored into blue. (k) has the higher min-value which only allows the higher speed values to render, indicating the high speed near the storm’s eye. (l) rotated visualization allowing a clearer view into the storm’s lower layers.

7 COMPARISON WITH OTHER MODALITIES

Apart from RadViz, there are other multivariate visualization interfaces that maintain the relationships of the channels and the samples. An obvious choice here is the method of Parallel Coordinates (PC) [47] which has been used by Wang et al. [25] and Guo et al. [24] aided by additional interactive displays (see Section 3). As argued in Section 3 we chose RadViz over PC mainly because of its ease of interaction. In the following, we present a comparison of PC with our RadViz-based MTE interface with regards to the faithfulness of the colorization that can be achieved.

To provide a fair comparison we used a PC implementation with similar capabilities than our RadViz-based MTE, which essentially is the vanilla implementation. For each axis we assigned a hue equivalent to the respective channel node in the MTE. The saturation decreased with the data point’s channel value just like in the MTE along the radial line toward the RadViz center. A sample’s RGB color is then the sum of the RGB colors interpolated on each axis, subsequently normalized. Figure 17(e) shows the colored PC display for the Hard X-Ray dataset (see Section 6.2).

Figure 17(a) shows the colorization results obtained with our system while Figure 17(b) shows the equivalent result obtained with the PC implementation. We observe significant differences with regards to fine detail resolution. The PC-based colorization appears more blurry and washed out, some detail is missing, and most importantly it cannot resolve the 4th phase (circled in Figure 14(a)). It essentially uses the same or very similar colors than for the 3rd phase.

Figure 17(d) plots the PC-colored points into the RadViz display. It can be easily noticed that they do not have the level of brightness uniformity and hue separability than the points in the MTE-based representation (Figure 17(c)). This is probably because the summation process in the PC does not provide the proper constraints for this. Conversely, by mapping the points directly into the HSL map their hues are more distinguishable and their brightness is similar.

7.1 Parallel Coordinates as a Free-Form Lens

For study purposes we also implemented an additional workflow where we used the parallel coordinates interface to brush data points by channel value, and then mapped these vectors into the MTE for colorization. Only the points selected appeared in the MTE and so these brushing operations essentially acted as a free-form lens. In addition, we also added a vector magnitude axis which served the role of the two magnitude sliders in Figure 1(E).

Akin to what has been reported in Bertini et al. [26], the interactions are more tedious since the user now needs to modify each axis to define a subset of data points and also the neighborhood relations are not as easily seen than on the RadViz display. But we found that it does allow for a more precise definition of neighborhoods in the MTE than what can be achieved with the circular lenses, and the scientists appreciated these capabilities as a verification tool.

It should be noted, however, that these more general regions can also be defined with our polygonal brushing tool (see Fig. 14(c)) although the colors assigned to these

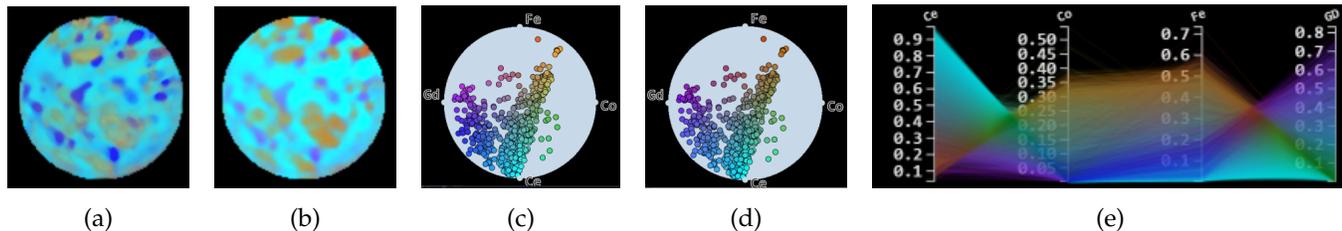


Fig. 17. Comparing the use of RadViz with Parallel Coordinates (PC) for data-driven colorization tasks, using the Hard X-Ray dataset. Rendering with (a) our RadViz-based MTE and (b) PC – the colors appear blotchy and the 4th phase is not differentiated (also refer to Figure 14(a)). (c-d) A RadViz-projection of the data points colored with (c) our MTE and (d) PC – the points are less differentiated by hue and vary in brightness. This explains the colorization deficiencies of PC. (e) PC display – each axis is assigned to a color ramp where the primary color is the same than for the respective channel node in the MTE.

regions are uniform, similar to those by Wang et al. [25] and Guo et al. [24].

8 EVALUATION

It became clear that the scientists we worked with all preferred the following workflow. They would begin with an automated colorization to produce volume renderings that would give a good overview of possible interesting phenomena, such as material phase changes. The next step commonly was to use lensing and clipping to isolate certain structures in the volume. These activities seemed to help the domain experts to learn about the distributions in information space and link them to their spatial renditions, using the color as a bridge. Finally, they would switch to the brushing interface and tag the spatial discoveries they made in solid color. This workflow follows the design goals DG1-DG4 presented in Section 4 in order.

8.1 Questionnaire and User Feedback

During one of the final sessions we asked each of the two teams for feedback on the following questionnaire:

Q1: Does our visualization tool help you in studying the composition of multivariate chemical data?

Q2: Would you consider using our tool in your work?

Q3: If yes, what do you feel is different when using our tool as compared to the existing tool you use for the same work?

Q4: Are there any significant changes or additional functionalities you would like to see in our tool in the future?

For both Q1 and Q2 both teams responded with a resounding yes. As for Q3, their current practice was to look at the different channels as separate volumes and examine them one by one or side by side. This made it difficult to notice the subtle interplay of all of the chemicals in 3D space. With our tool, on the other hand, they could fluidly examine the multi-material data in one picture, without needing to tediously switch from channel to channel. One of the teams specifically mentioned that they thought of our interface as a one-stop-shopping experience. They could pursue initial hunches, then refine, accept, or reject them, and if there was merit they could go all the way to full documentation of these findings, using our manual colorization interface.

All scientists stated without hesitation that our tool can save them a tremendous amount of time and effort. While they could find most of the structures reported in the paper before, it took them a long time to do so, on the order of days and even weeks. Our tool allowed them to find these same structures within 10-15 minutes, and in a publishable

format to boot. Essentially, our tool significantly accelerates time to discover which in consequence has the propensity to accelerate science, achieving progress at a faster speed. In addition, we also observed that, afforded by our streamlined interface, there is also a potential for new discoveries. We saw some onsets of this already in our evaluation sessions (for example, the discovery of a 4th phase, see Section 6.2).

When asked what features they would like to see implemented (Q4) one of the scientists said that he would welcome a histogram augmented with a range selection slider for each channel. Since this is essentially a 1D transfer function we asked if he would prefer this selection interface over the MTE. He replied that he would see it mainly as a complementary feature for additional information. To address this need, we re-purposed the parallel coordinate interface as discussed in Section 7.1.

9 DISCUSSION AND LIMITATIONS

Our method shares a common problem of post-classified semi-transparent volume rendering where an interpolated scalar sample receives a color via a transfer function's color map which is subsequently blended with another sample's color along the ray. Assume there are two materials and the colors assigned are blue and red; mixing these two colors in semi-transparent rendering mode results in the (secondary) color purple. Similarly, yellow and red color mix to orange. If a third material were assigned the color purple or the color orange ambiguities could arise. These effects are well known and several solutions have been proposed (see [48]), and we plan to incorporate one of these into a future implementation of our software. For now, user interactions we provide, such as volume rotation, slicing, and value range selection as well as MTE lensing and rotation can help resolve ambiguities of this nature if they occur.

For some of the images the regional boundaries appear somewhat fluid and unsharp. This is due to the continuous color mapping in the MTE which assigns smoothly changing colors to adjacent volume points in areas where the material composition gradually changes. It is a consequence of the partial volume effect which is further amplified in our case since a volume point is a vector of several materials and not just a scalar density. In contrast, transfer functions used in conventional volume rendering often use uniform solid colors (with perhaps some fading at the density-range edges) for a range of density values. The approaches of Wang et al. [25] and Guo et al. [24] also make these kinds of uniform color assignments to segmented regions in the data

space, and as a consequence these displays appear sharper. Notably, our polygonal brushing tool achieves similar effects especially when the polygonal regions have some gaps between them, as can be seen in Fig. 14c.

We currently use opacity only in a uniform way as a means to allow the visualization of structures deeper in the volume via making the outer volume layers more transparent. This turned out to be a helpful function in the discovery of hidden interior structures, as illustrated in Fig. 13. A more sophisticated opacity-driven enhancements might associate a lens with an opacity slider such that the visibility of the points covered by the lens could be more specifically controlled. Or one might associate opacity transfer functions with the axes of the proposed parallel coordinate-based interaction tool. This would not be difficult to implement and we plan this as a future extension.

For the coupled MTE-PC display one could imagine a bidirectional workflow in which a user would first use a circular lens in the RadViz display and then refine these vectors further in the PC interface, and vice versa. It would unify the strengths of both displays and would be a novel application of the two coordinated displays not mentioned by Bertini et al. In this scheme, the RadViz-based lens would capture groups of similar points and the PC lens would refine this group. The volume rendering would then provide real-time feedback for both of these lenses and support the user in refining the volumetric structures.

Finally, a future version will also include an option to set the background of the volume rendering to either black or white. Studies have shown that users prefer a black background for its aesthetics, but also believe that a white background makes small detail more discriminable [49].

10 CONCLUSIONS

We have presented a methodology and tool that allows the visualization of multivariate (multi-channel) volume data with full awareness of the contributions of the individual channel attributes. We achieve this by fusing a traditional GPU-accelerated volume renderer with an information visualization front-end that allows users to assign colors and opacities to voxel vectors according to their multivariate value distributions and attribute relations. The information visualization interface, fused with a colormap and augmented by various interactive filtering facilities, links the statistical characteristics of the multivariate volume data with their appearance characteristics in the spatial domain.

A particular strength of our tool is that by fusing complex volumetric multivariate data into a single display the time to discovery can be significantly sped up. This in consequence allows science to progress at a faster speed. In fact, our tool has already led to the discovery of potentially new insight, facilitated by its ease of use.

Even though we have evaluated our tool mainly in material science, we believe that it can be applied in any other domain that has challenges in the analysis of multivariate volumetric data. Especially suitable are fields that acquire volumetric data by ways of multi-valued or multi-modal scanning, such as industrial CT, medicine, or bio-science. We have already presented an application for the latter in Fig. 11. But even atmospheric science or meteorology (see Fig.

16) which acquire or simulate multi-valued data on large scale grids could benefit from our general methodology. We plan to study these types of data in the future.

ACKNOWLEDGMENTS

This research was partially supported by NSF grants IIS 1527200 and 1941613 and by Brookhaven National Laboratory LDRD grant 16-041. Huolin L. Xin was supported by NSF Award CHE-1900401. This research used resources of the hard x-ray nanoprobe beamline (HXN) at 3ID of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704.

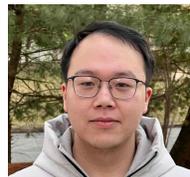
REFERENCES

- [1] R. Drebin, L. Carpenter, and P. Hanrahan, "Volume rendering," *ACM Siggraph Computer Graphics*, vol. 22, no. 4, pp. 65–74, 1988.
- [2] M. Levoy, "Display of surfaces from volume data," *IEEE Computer graphics and Applications*, vol. 8, no. 3, pp. 29–37, 1988.
- [3] J. Kniss, G. Kindlmann, and C. Hansen, "Interactive volume rendering using multi-dimensional transfer functions and direct manipulation widgets," in *IEEE Visualization*, 2001, pp. 255–262.
- [4] J.-N. Zhang et al., "Trace doping of multiple elements enables stable battery cycling of licoo 2 at 4.6 v," *Nature Energy*, vol. 4, no. 7, pp. 594–603, 2019.
- [5] W. Li et al., "Dynamic behaviour of interphases and its implication on high-energy-density cathode materials in lithium-ion batteries," *Nature communications*, vol. 8, no. 1, pp. 1–10, 2017.
- [6] J. Steiner et al., "Targeted surface doping with reversible local environment improves oxygen stability at the electrochemical interfaces of nickel-rich cathode materials," *ACS Applied Materials & Interfaces*, vol. 11, no. 41, pp. 37885–37891, 2019.
- [7] I. Markus, F. Lin, K. Kam, M. Asta, and M. Doeff, "Computational and experimental investigation of ti substitution in li1 (ni x mn x co1-2 x-y ti y) o2 for lithium ion batteries," *Journal of Physical Chemistry Letters*, vol. 5, no. 21, pp. 3649–3655, 2014.
- [8] W. Harris et al., "Characterization of 3d interconnected microstructural network in mixed ionic and electronic conducting ceramic composites," *Nanoscale*, vol. 6, no. 9, pp. 4480–4485, 2014.
- [9] R. Hovden and D. Muller, "Electron tomography for functional nanomaterials," *MRS Bulletin*, vol. 45, no. 4, p. 298–304, 2020.
- [10] M. Hanwell et al., "Tomviz: Open source platform connecting image processing pipelines to gpu accelerated 3d visualization," *Microscopy and Microanalysis*, vol. 25, no. S2, p. 408–409, 2019.
- [11] H. Yan, P. Voorhees, and H. Xin, "Nanoscale x-ray and electron tomography," *MRS Bulletin*, vol. 45, no. 4, p. 264–271, 2020.
- [12] P. Ljung et al., "State of the art in transfer functions for direct volume rendering," in *Computer Graphics Forum*, vol. 35, no. 3, 2016, pp. 669–691.
- [13] A. Kaufman and K. Mueller, "Overview of volume rendering," *The Visualization Handbook*, vol. 7, pp. 127–174.
- [14] U. Ayachit, *The Paraview Guide: A Parallel Visualization Application*. Kitware, Inc., 2015.
- [15] C. ImageVis3D, "An interactive visualization software system for large-scale volume data," 2015.
- [16] R. Maciejewski, I. Woo, W. Chen, and D. Ebert, "Structuring feature space: A non-parametric method for volumetric transfer function generation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1473–1480, 2009.
- [17] F.-Y. Tzeng and K.-L. Ma, "A cluster-space visual interface for arbitrary dimensional classification of volume data," in *Proc EG-IEEE TCVG Conference on Visualization*, 2004, pp. 17–24.
- [18] L. Wang, X. Zhao, and A. Kaufman, "Modified dendrogram of attribute space for multidimensional transfer function design," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 1, pp. 121–131, 2012.
- [19] J. M. Kniss, R. Van Uitert, A. Stephens, G.-S. Li, T. Tasdizen, and C. Hansen, "Statistically quantitative volume visualization," in *IEEE Visualization*, 2005, pp. 287–294.

- [20] M. Haidacher, S. Bruckner, A. Kanitsar, and M. E. Gröller, "Information-based transfer functions for multimodal visualization." in *VCBM*, 2008, pp. 101–108.
- [21] H. S. Kim, J. Schulze, A. Cone, G. Sosinsky, and M. Martone, "Dimensionality reduction on multi-dimensional transfer functions for multi-channel volume data sets," *Information Visualization*, vol. 9, no. 3, pp. 167–180, 2010.
- [22] A. Abbasloo, V. Wiens, M. Hermann, and T. Schultz, "Visualizing tensor normal distributions at multiple levels of detail," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 975–984, 2016.
- [23] E. Mörth, I. Haldorsen, S. Bruckner, and N. Smit, "Paraglyder: Probe-driven interactive visual analysis for multiparametric medical imaging data," in *Proc. Computer Graphics International*, 2020, pp. 351–363.
- [24] H. Guo, H. Xiao, and X. Yuan, "Multi-dimensional transfer function design based on flexible dimension projection embedded in parallel coordinates," in *IEEE PacificVis*, 2011, pp. 19–26.
- [25] J. Wang, X. Liu, H.-W. Shen, and G. Lin, "Multi-resolution climate ensemble parameter analysis with nested parallel coordinates plots," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 81–90, 2016.
- [26] E. Bertini, L. Dell'Aquila, and G. Santucci, "Springview: Cooperation of radviz and parallel coordinates for view optimization and clutter reduction," in *IEEE Coordinated and Multiple Views in Exploratory Visualization (CMV)*, 2005, pp. 22–29.
- [27] P. Hoffman, G. Grinstein, K. Marx, I. Grosse, and E. Stanley, "Dna visual and analytic data mining," in *Proc IEEE Visualization*, 1997, pp. 437–441.
- [28] T. Pfaffelmoser, M. Reitingner, and R. Westermann, "Visualizing the positional and geometrical variability of isosurfaces in uncertain scalar fields," in *Computer Graphics Forum*, vol. 30, no. 3. Wiley Online Library, 2011, pp. 951–960.
- [29] C. Lundström, P. Ljung, and A. Ynnerman, "Multi-dimensional transfer function design using sorted histograms," in *EG/IEEE VGTC Workshop on Volume Graphics*, 2006.
- [30] L. Wang and A. Kaufman, "Importance driven automatic color design for direct volume rendering," in *Computer Graphics Forum*, vol. 31, no. 3pt4, 2012, pp. 1305–1314.
- [31] R. Bramon, M. Ruiz, A. Bardera, I. Boada, M. Feixas, and M. Sbert, "Information theory-based automatic multimodal transfer function design," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 4, pp. 870–880, 2013.
- [32] P. Sereda, A. Vilanova, and F. Gerritsen, "Automating transfer function design for volume rendering using hierarchical clustering of material boundaries." in *EuroVis*, 2006, pp. 243–250.
- [33] C. Y. Ip, A. Varshney, and J. JaJa, "Hierarchical exploration of volumes using multilevel segmentation of the intensity-gradient histograms," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2355–2363, 2012.
- [34] S. Liu, B. Wang, J. Thiagarajan, P.-T. Bremer, and V. Pascucci, "Multivariate volume visualization through dynamic projections," in *IEEE Symp. Large Data Analysis & Vis. (LDAV)*, 2014, pp. 35–42.
- [35] F.-Y. Tzeng, E. B. Lum, and K.-L. Ma, "An intelligent system approach to higher-dimensional classification of volume data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 3, pp. 273–284, 2005.
- [36] F. de Moura Pinto and C. M. Freitas, "Design of multi-dimensional transfer functions using dimensional reduction," in *Proc. EG/IEEE VGTC Conference on Visualization*, 2007, pp. 131–138.
- [37] S. Muraki, T. Nakai, Y. Kita, and K. Tsuda, "An attempt for coloring multichannel mr imaging data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 7, no. 3, pp. 265–274, 2001.
- [38] A. Lu and D. S. Ebert, "Example-based volume illustrations," in *IEEE Visualization*, 2005, pp. 655–662.
- [39] G. Kindlmann and D. Weinstein, "Hue-balls and lit-tensors for direct volume rendering of diffusion tensor fields," in *IEEE Visualization*, 1999, pp. 183–524.
- [40] G. Kindlmann, D. Weinstein, and D. Hart, "Strategies for direct volume rendering of diffusion tensor fields," *IEEE Transactions on Visualization and Computer Graphics*, vol. 6, no. 2, pp. 124–138, 2000.
- [41] T. Welsh, M. Ashikhmin, and K. Mueller, "Transferring color to greyscale images," in *ACM Transactions on Graphics*, vol. 21, no. 3, 2002, pp. 277–280.
- [42] S. Cheng, W. Xu, and K. Mueller, "Colormap nd: A data-driven approach and tool for mapping multivariate data to color," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 2, pp. 1361–1377, 2019.
- [43] —, "Radviz deluxe: An attribute-aware display for multivariate data," *Processes*, vol. 5, no. 4, p. 75, 2017.
- [44] G. Ferreira *et al.*, "Micro x-ray fluorescence elemental analysis of spheroid cultures of human neuroblastoma cells," *Radiation Physics and Chemistry*, vol. 167, p. 108190, 2020.
- [45] Y. Shao *et al.*, "Making li-air batteries rechargeable: Material challenges," *Adv. Functional Materials*, vol. 23, no. 8, pp. 987–1004, 2013.
- [46] H. Yan *et al.*, "Multimodal hard x-ray imaging with resolution approaching 10 nm for studies in material science," *Nano Futures*, vol. 2, no. 1, p. 011001, 2018.
- [47] A. Inselberg and B. Dimsdale, "Parallel coordinates: a tool for visualizing multi-dimensional geometry," in *Proc. IEEE Visualization*, 1990, pp. 361–378.
- [48] L. Kühne, J. Giesen, Z. Zhang, S. Ha, and K. Mueller, "A data-driven approach to hue-preserving color-blending," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2122–2129, 2012.
- [49] J. Giesen, K. Mueller, E. Schuberth, L. Wang, and P. Zolliker, "Conjoint analysis to measure the perceived quality in volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1664–1671, 2007.



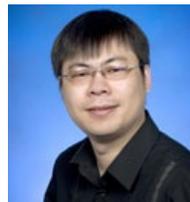
Ayush Kumar has a Ph.D. degree in Computer Science from Stony Brook University. His research interests include Machine Learning, Visual Analytics, and Eye Tracking with a special focus on Text and spatio-temporal data. He is now a Senior Member of Technical Staff at ATT Labs, a Fellow at Schepens Eye Research Institute, Harvard Medical School, and a Research Assistant Professor at the Computer Science Department of Stony Brook University.



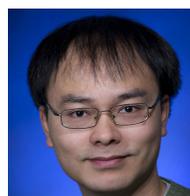
Xinyu Zhang earned his B.E. in Computer Science from Shandong University in 2019. Currently, he is a PhD candidate at Stony Brook University in Computer Science. His research interests include multivariate data analysis, scientific visualization, and reinforcement learning.



Huolin L. Xin received his Ph.D. in Physics from Cornell University in 2011. He is now a tenured full professor in the Department of Physics and Astronomy at the University of California, Irvine. His expertise is advanced analytical imaging for materials and chemical physics.



Hanfei Yan earned his B.S. in Physics in 1999 from Peking University, a M.S. in Physics from U Mass Amherst in 2001, and Ph.D. from Columbia University in 2005. After graduation, he was a postdoc at the Center for Nanoscale Materials of Argonne National Laboratory to develop x-ray nanofocusing optics. In 2007, he joined Brookhaven National Laboratory and currently is the lead beamline scientist of the hard x-ray nanoprobe beamline of the National Synchrotron Light Source II.



Xiaojing Huang earned his B.S. in Physics, Fudan University, China, 2003, and an M.S. and Ph.D. in Physics from Stony Brook University, Stony Brook, NY, in 2005 and 2009, respectively. He accepted a postdoc position at the NSLS-II to develop x-ray nanofocusing at Brookhaven National Laboratory (BNL). Currently, he is Physicist at NSLS-II, BNL.



Wei Xu (M'2015) received the BS and MS degrees in computer science from Zhejiang University, in 2004 and 2006, and the PhD degree in computer science from Stony Brook University in 2012. She is currently a Computer Scientist at the Computational Science Initiative of Brookhaven National Laboratory and a visiting Research Assistant Professor at the Computer Science Department of Stony Brook University.



Klaus Mueller has a PhD in computer science and is currently a professor of computer science at Stony Brook University and a senior scientist at Brookhaven National Lab. His current research interests include explainable AI, visual analytics, data science, and medical imaging. He won the US National Science Foundation Early CAREER Award, the SUNY Chancellor's Award for Excellence in Scholarship & Creative Activity, and the IEEE CS Meritorious Service Certificate. His 300+ papers were cited over 12,500 times.