(2008 IEEE Visualization Design Contest Winner) Linking Multi-dimensional Feature Space Cluster Visualization to Surface Extraction from Multi-field Volume Data

Lars Linsen   Tran Van Long   Paul Rosenthal

School of Science and Engineering
Jacobs University, Bremen, Germany
E-mail: {l.linsen, v.tran, p.rosenthal}@jacobs-university.de

Data sets resulting from physical simulations typically contain a multitude of physical variables. It is, therefore, desirable that visualization methods take into account the entire multi-field volume data rather than concentrating on one variable. We present a visualization approach based on surface extraction from multi-field volume data. The extracted surfaces segment the data with respect to an underlying multi-variate function. Decisions on segmentation properties are based on the analysis of a multi-dimensional feature space. The feature space exploration is performed using an automated multi-dimensional hierarchical clustering method. The hierarchical clusters are shown as a cluster tree in a 2D radial layout. In the cluster tree layout, the user can select clusters of interest. A selected cluster in feature space corresponds to a segmenting surface in object space. Based on the segmentation property induced by the cluster membership, we extract surfaces from the volume data.

This approach was applied to the 2008 IEEE visualization design contest data set. The provided time-varying data set is a simulation of the propagation of an ionization front instability. It includes a variety of different attributes, including density, temperature, mass abundances of eight chemical species, and velocity. The size of each of the 200 time steps is $37 \cdot 10^6$ points resulting in 1.7 GB of data per time step. For performance purposes, we decided to spatially sample the data. As no numerical details were available on how the simulation was produced, the best decision towards an unbiased sampling was to sample randomly. As a result, we obtained a set of uniform randomly distributed points in the volumetric space, where each point carries multiple properties. We experienced that one million points per time step were sufficient to robustly reproduce clustering results. We directly extract surfaces from the data by computing individual points on the surface, supported by an efficient neighborhood computation. The extracted surface points are rendered using point-based rendering operations. Our approach combines methods in scientific (spatial) visualization for object-space operations with methods in information (non-spatial) visualization for feature-space operations.

HIERARCHICAL DENSITY-BASED CLUSTERING

When dealing with large data sets with many records, clustering has proven to be extremely useful. Clustering is a method to partition a data set into subsets of similar observations. Each subset is called a *cluster*, which consists of observations that are similar within themselves and dissimilar to observations of other clusters. Cluster analysis tasks for multidimensional data include finding areas where individual records group together to form a cluster.

We analyzed the data by looking into the density of the distribution of the multi-dimensional feature space. Clusters are formed by areas of locally high density. Density computations have to be performed in the original high-dimensional space, as projection to lower-dimensional space imposes assumptions and thus distorts the data. Our density computations are based on subdividing the domain into hypercubes and evaluating density functions over the spatial subdivision. The clustering is done hierarchically by iteratively splitting each density cluster into subclusters (if any). Any splitting of a cluster into multiple higher-density sub-clusters is recorded [1].

CLUSTER TREE VISUALIZATION

To display the hierarchical structure of the density clusters, we generate a hierarchical density cluster tree and show it using a 2*D* radial layout. The root of the tree is placed in the origin of a unit circle, while the leaves are evenly distributed on the circle. The internal nodes are placed on concentric circles corresponding to its depth. The clusters are visually represented as colored disks, whose radii encode the size of the clusters and whose colors encode the position within the radial layout using the *HSV* color space, see Figure 1 (lower left).

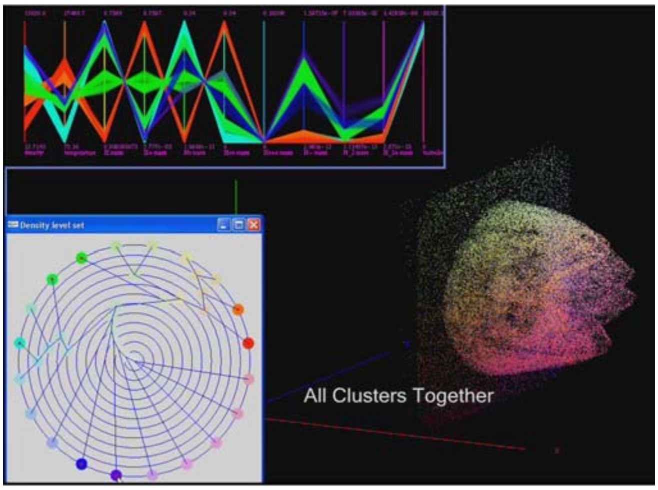

Figure 1: Linked views of the hierarchical cluster tree visualization (lower left) to parallel coordinates (upper) and object-space rendering (lower right). The automatically detected hierarchical clusters are interactively selected in the cluster tree visualization and simultaneously displayed in the respective linked views.

LINKED VIEWS

The cluster tree visualization supports the comprehension of the structure of clusters within the multidimensional feature space. To also support the visualization of the value ranges for the individual dimensions, we employ parallel coordinates. Linked views are used to select clusters in the cluster tree representation and display them in a parallel coordinates representation, see Figure 1 (upper). Colors are used for identification. The applied coloring scheme assures unique color coding. In Figure 1, several clusters have been selected simultaneously. In the radial layout, they are represented by disks colored with higher saturation and exhibiting a small red dot in its center. In parallel coordinates, the values of the selected clusters with respect to the multi-variate field are visualized.

In addition, we support another linked view that renders the spatial distribution of clusters in the volumetric object space. Figure 1 (lower right) shows the volumetric distribution of the points belonging to the selected clusters. The user interacts with the cluster visualization and gets multiple linked views on the data.

SURFACE EXTRACTION AND RENDERING

The multidimensional cluster membership induces a segmentation in object space. Based on this property, we can extract a surface from the volume data using direct surface extraction [3]. More precisely, we extract the isosurface with isovalue 0.5 to the characteristic function of the cluster. We directly extract our surfaces without prior resampling or grid generation. The surface extraction computes individual points on the surface, which is supported by a neighborhood computation using *k*d-trees and an efficient indexing scheme.

The extracted surface points are rendered using point-based rendering operations. An interactive approach is provided using image-based point cloud rendering [4]. The lit surface points are directly rendered to the screen. Holes are filled with image-based filters to produce hole-free surface renderings, see Figure 2(c).

A second approach for rendering the surface points is splat-based ray tracing [2], producing renderings with global illumination but not at interactive rates. For each surface point a circular splat is fit to the surface using a least-squares approach providing also a normal map of the splat. The resulting splats are ray traced using the normal map, see Figure 2(a).

SCIENTIFIC QUESTIONS

Symmetry of the Shadow Instability. First we want to clarify the three-dimensional shape of the observed shadow instability. For this purpose, we randomly resampled the data set of time step 90. Afterwards, it was clustered in feature space according to all attributes. We extracted the surface segmenting the cluster of the ambient gas with a temperature around 72 K from the shocked and ionized gas. This surface represents the front of the ionization process. A ray-traced picture of the surface is shown in Figure 2(a).

Using the same view as in Figure 1, it is clearly visible that the shadow instability is not symmetric around the *x*-axis. Instead it is symmetric with respect to the y = 124 and the z = 124 planes. Furthermore, it is symmetric with respect to the two planes that are perpendicular to the *x*-axis and lie diagonal in the *yz*- plane. The symmetry around the *x*-axis is mainly broken by the eight “fingers” representing the instability front.

Prevalence of H2. We chose time step 99 to show the regions, where H2 is most prevalent. The randomly resampled data set was again clustered in feature space. In Figure 2(b) we show a point rendering of two important clusters. The points of the cluster with most H2 are rendered in green. To have a context for this cluster, we additionally render the points of the cluster with highest turbulence in red. These points indicate the zone of the shocked gas in the ionization process.

The picture shows that most of the H2 is generated at the very beginning of the ionization process. It is most prevalent directly at the beginning of the zone with shocked gas and high turbulence. The properties of the extracted clusters - more precisely, the average attribute values - also show that in the regions where H2 is most prevalent, most of He is already ionized to He+ but only some H+ was ionized from H.

Radiation Fingers. To see how thick the first fingers of radiation are, that break through the front, we have to study data early in time. We took time step 10, where this phenomenon first clearly appears. Again, a non-equidistant sampling followed by a multi-dimensional clustering was applied. Thereafter, two segmenting surfaces were extracted and rendered using image-space point cloud rendering, see Figure 2(c).

The bluish surface segments the ambient gas from the rest and gives the context for a better understanding of the picture. The red surface segments the first finger of radiation breaking through. The picture was rendered with a viewport parallel to the *xy*-plane to allow a better judgment of the size. The diameter of the finger of radiation is approximately 0.017 parsecs.

Cause of Turbulence. To answer the question if turbulence is stirred up in the front of the shadow instability, we again refer to Figure 2(b). As mentioned in Section , time step 99 was resampled and clustered in feature space. To have an indication for turbulence, we introduced the magnitude of the curl for the provided velocity field as an additional data dimension. The points of the cluster with high turbulence are rendered in red.

It is clearly visible, that most turbulence is present in the region of shocked gas, i. e., directly behind the ionization front. So we have a real strong indication that the shadow instability stirs up the turbulence in the gas.

H2 Formation and Turbulence. To answer the last two key questions we refer to Figures 2(d) and (e). The data set for time step 99 was resampled non-equidistantly and clustered in feature space. Afterwards, three clusters were extracted. The points of these clusters were rendered with different colors to provide a three-dimensional view of the cluster distribution.

The points of the cluster with highest H2 are colored green. All red and blue points represent points of high turbulence, cf. Figure 2(b). These points were split into two clusters representing points with low H+ density, colored red, and points with high H+ density, colored blue.

Most of H2 is generated at the very beginning of the turbulence cluster, where H+ density is low, i.e., where it is lower than 0.2. This indicates that turbulence is essential for the formation of H2, but the needed electrons do mostly not come from H+. Instead, a high density of He+, i.e., greater than 0.15, is observable in the regions of H2 formation. This also heavily correlates with the presence of H− and H2+.

More precisely, one gets the following impression from the procedures causing the turbulence: When the ionization front reaches the ambient gas with $\rho H \approx 0.76$ and $\rho He \approx 0.24$, both chemical species are ionized at approximately the same speed. However, a high amount of H+ is converted into H2 with the help of the free electrons from He+. This process, and probably also the break-up of H2, afterwards lead to the observable turbulences behind the H2 front. This behavior can be observed at the front of the instability as well as at the front of the “normal” ionization front.

CONCLUSION

Multi-dimensional clustering in feature space can be a powerful tool when appropriate interactive mechanisms are provided to explore the clusters. Our cluster tree visualization supports an intuitive understanding of the data distribution in feature space. When linked to parallel coordinates, the range of values of each cluster in the individual dimensions can be investigated. This linked view proved extremely useful for answering question from the application domain and for validating hypotheses. When linked to object space visualization, the spatial location of the detected clusters of interest can be observed and analyzed. We provided a visualization of cluster points in object space as well as direct surface extraction from the randomly distributed sample points. The surfaces represent the boundary of an object-space segmentation induced by the feature space clustering. We extract them in a point cloud surface representation and use point-based rendering techniques for their display. With this linked view set-up, we were able to answer questions that involve both object- and feature-space investigations.

REFERENCES

[1] L. Linsen, T. V. Long, P. Rosenthal, and S. Rosswog. Surface extraction from multi-field particle volume data using multi-dimensional cluster visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1483–1490, Nov.-Dec. 2008.

[2] L. Linsen, K. Müller, and P. Rosenthal. Splat-based ray tracing of point clouds. *Journal of WSCG*, 15(1–3), 51–58, 2007.

[3] P. Rosenthal and L. Linsen. Direct isosurface extraction from scattered volume data. In *Proceedings of Eurographics/IEEE-VGTC Symposium on Visualization*, pages 99–106, 2006.

[4] P. Rosenthal and L. Linsen. Image-space point cloud rendering. In *Proceedings of Computer Graphics International*, pages 136–143, 2008.

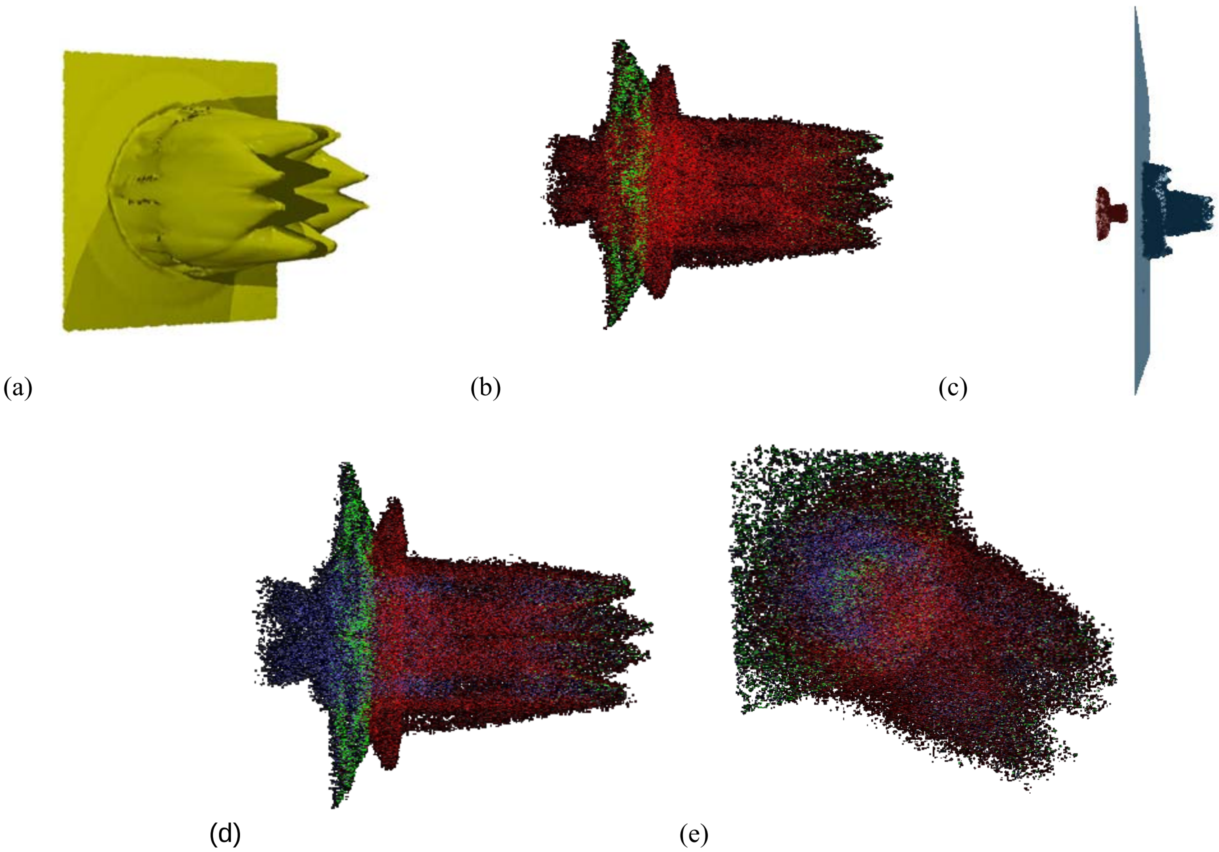

Figure 2: Results of the interactive visual exploration of the ionization front instability propagation.