

# Segmentation and Visualization of Multivariate Features using Feature-Local Distributions

Kenny Gruchalla<sup>1,2</sup>, Mark Rast<sup>2,3</sup>, Elizabeth Bradley<sup>2</sup>, and Pablo Mininni<sup>4,3</sup>

<sup>1</sup> National Renewable Energy Laboratory, Golden, Colorado

<sup>2</sup> University of Colorado, Boulder, Colorado

<sup>3</sup> National Center for Atmospheric Research, Boulder, Colorado

<sup>4</sup> Universidad de Buenos Aires, Argentina

**Abstract.** We introduce an iterative feature-based transfer function design that extracts and systematically incorporates multivariate feature-local statistics into a texture-based volume rendering process. We argue that an interactive multivariate feature-local approach is advantageous when investigating ill-defined features, because it provides a physically meaningful, quantitatively rich environment within which to examine the sensitivity of the structure properties to the identification parameters. We demonstrate the efficacy of this approach by applying it to vortical structures in Taylor-Green turbulence. Our approach identified the existence of two distinct structure populations in these data, which cannot be isolated or distinguished via traditional transfer functions based on global distributions.

## 1 Introduction

We describe an iterative analysis and visualization technique that allows domain experts to interactively segment, group, and investigate individual multivariate volumetric structures by incorporating *feature-local* statistics into a multivariate transfer function. The opacity component of a volume-rendering transfer function is used as an initial threshold to create a binary segmentation of the data volume. Individual structures are then identified through a connected-component analysis, and variable distributions are calculated in the spatial extents of the individual structures. Through an interactive table, structures can be filtered and selected based on their local statistical properties (e.g., central moments of the feature-local distributions), which are then added to the volume-rendering transfer function as an additional dimension. Users can selectively iterate various stages of this process, allowing them to progressively refine the visualization, improve their understanding of the multivariate properties unique to each structure, and investigate the correlations between variables across localities at multiple scales.

As a proof of concept, this tool was applied to data from a simulation of forced Taylor-Green turbulence [1]. Analysis of the high-vorticity structures in the Taylor-Green turbulence data volume by the means described in this paper

revealed distinct vortical structure populations with quite different flow properties. These populations cannot be isolated or distinguished using a traditional multidimensional transfer function based on the global distributions of flow quantities.

## 2 Related Work

In general, the goal of feature-based visualization is to extract physically meaningful structures from the data, showing only those *features* that are of interest to the researcher. Feature-based visualization of turbulent flow typically targets physical characteristics of features using either image processing [2–5] or topological analysis [6–8]. Post et al. [9] have provided an extensive survey of feature-based visualization techniques as applied to flow data. Our extraction approach is based on an image-processing technique, where the physical characteristics of interest are defined by a multidimensional opacity function and the resulting connected opaque structures are extracted using a connected-component labelling algorithm [10].

There is considerable literature in the field of multivariate visualization. A comprehensive review is beyond the scope of this paper. Wong and Bergeron [11] provide a survey of techniques appropriate for the visualization of abstract multivariate data. Bürger and Hauser [12] provide a survey of visualization techniques specifically for multivariate scientific data, with an emphasis on volumetric data. One common approach is to visualize the relationships or correlations between multiple variables, as our technique does. Local statistical complexity has been used to identify regions of multivariate data that communicate large amounts of information [13]. Multi-field graphs, which give a useful overview of correlation fields between variables, can help guide the selection of promising correlations [14]. Multidimensional transfer functions were introduced to investigate and exploit these kinds of correlations [15]. Park et al. [16] applied this in the context of turbulence data, incorporating flow field properties such as velocity, curl, helicity, and divergence into multidimensional transfer functions to visualize features in flow fields. Our incorporation of multiple fields is different from these approaches in several ways. We provide the user the ability to interactively investigate *local* multivariate relationships. The locality can be constrained to the currently visualized space or applied systematically to each individual connected-feature. These localities are defined by the opacity component of an initial transfer function, which can support multiple dimensions. The result is a system that can investigate and visualize the multivariate relationships between a spectrum of scales, which is particularly important when investigating phenomena believed to operate across multiple scales, such as turbulence.

Statistical guides and the exploitation of various statistical properties are common elements in interactive volume visualization. The underlying principle of data-driven transfer function design is to use information derived from the data to guide the user in the design process. For example, all of the multidimensional transfer function techniques described above [15, 16] incorporate

global-data histograms and/or scatter plots to provide statistical context. The use of *localized* statistics is increasingly common in interactive volume visualization. The Contour Spectrum technique, which is used to determine isovalues in unstructured meshes, incorporates a variety of isosurface metrics (e.g., surface area, volume, and mean gradient magnitude) into the selection interface [17]. Tenginakai et al. [18] use localized higher order central moments based on a local kernel to find isosurface boundaries. Correa and Ma [19] use a multi-scale analysis to incorporate a derived scale field that represents the relative size of the local features at each voxel into a multidimensional transfer function. Lundström et al. [20] introduced a partial range histogram – computed locally, over a block-partition of space – to help distinguish tissue boundaries in medical data. Our use of local statistics is novel in two ways. First, the locality is based on the geometry of extracted structures, which is user-defined through a multidimensional transfer function. Second, we provide systematic access to the local statistical measures and to their distributions for each extracted structure. The user can focus the visualization and analysis by employing the central moments of these local distributions to sort, filter, and select the individual structures. In this way, we generalize the concept of spatial locality to include both structural and physical properties of the flow, thereby facilitating effective interactive feature definition, identification, and property extraction.

### 3 Methods and Implementation

To evaluate the utility of the techniques described in this paper in the analysis of turbulent structures, we have implemented them in the open-source VAPOR [21, 22] visualization and analysis environment, which is a widely deployed toolset in the turbulence community that specifically targets time-varying, multivariate, large-scale data sets. Our technique can be divided into four steps: *selection*, *clustering*, *attribute calculation*, and *quantization*.

The selection step is used to create a binary segmentation, denoting which of the voxels are isolated from the original data. We approach the selection step by thresholding, using the opacity contribution from a multidimensional transfer function. In addition to providing a facility to visualize correlations between multiple variables, multidimensional transfer functions are far more expressive than traditional one-dimensional transfer functions, allowing the extraction of features that have overlapping data values in one dimension [23]. To avoid the exponential memory requirements of the general  $n$ -dimensional approach, we decompose the transfer function into  $n \times m$  one-dimensional transfer functions, similar to the separable transfer function concept discussed by Kniss et al. [24]. We also handle the user interface issues of interacting with an  $n$ -dimensional transfer function via a decomposition approach, allowing the user to interact with either  $n$  one-dimensional interfaces or  $(n^2 - n)/2$  two-dimensional interfaces. Transfer function widgets are used to specify the function. Multiple widgets can be created by the user to isolate multiple regions in  $n$ -dimensional space. When

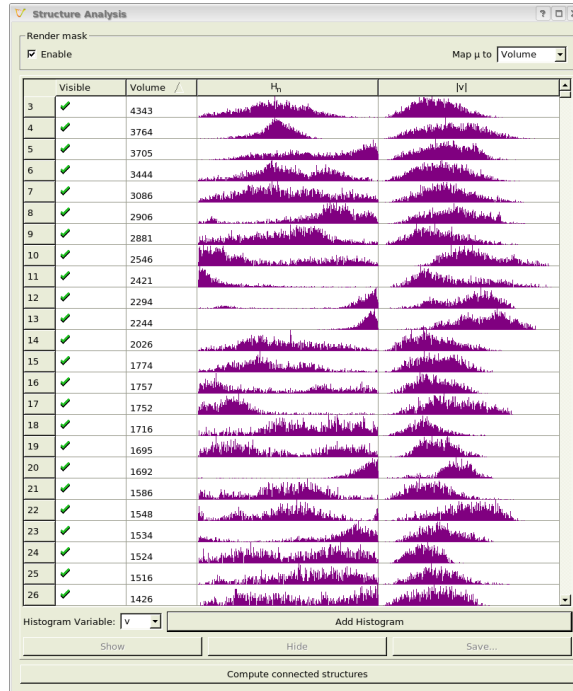
selected, widgets are visually linked across the decomposition, relaying each one’s contribution in  $n$ -dimensional space.

The clustering step classifies the selected voxels into coherent regions. We have chosen a connected-component labelling scheme as our clustering step. For the purposes of labelling, we treat our data as a binary volume, using the opacity function from our decomposed  $n$ -dimensional transfer function as a threshold to define voxels as transparent or opaque. The connected-component algorithm is then used to identify and uniquely label regions of connected opaque voxels.

In the next step, the attributes of the clustered regions are calculated. The attributes we are concerned with are feature-local statistics. i.e., data variable distributions bounded to the spatial extents of the individual clusters and their central moments. The user can interactively select field variables from the data set or import derived fields directly from external analysis packages, which can be linked to the data volume by metadata descriptors [21]. The distribution local to each structure is calculated on each of these fields, and the corresponding *feature-local histogram* is presented in an interactive table of structure attributes (see Figure 1). Subsets of interest can be interactively selected through sorting, filtering, and selecting the structures based on their attributes (e.g., their volume and the central moments of their local distributions). The quantization step incorporates this filtered data into the transfer function as an additional dimension, providing an interactive mechanism to explore the structures.

## 4 Results and Discussion

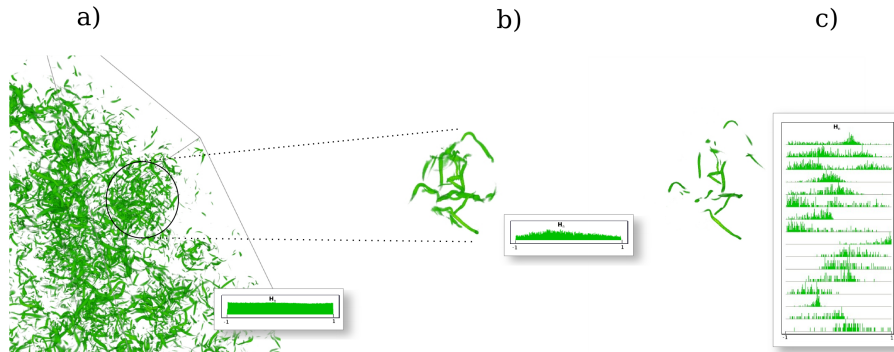
We demonstrate the power of this design by using it to analyze three-dimensional data from an incompressible forced Taylor-Green turbulence simulation [1]. We examined two flow properties: flow vorticity and helicity. Vorticity is the point-wise curl of the a velocity field,  $\boldsymbol{\omega} = \nabla \times \boldsymbol{v}$ . Helicity is the cosine of the angle between velocity and vorticity vectors, a scalar quantity,  $H_n = \frac{\boldsymbol{v} \cdot \boldsymbol{\omega}}{|\boldsymbol{v}| |\boldsymbol{\omega}|}$ . Relative to the size of the volume, these Taylor-Green data are characterized by many small vortical features, and when applied to the full volume, a threshold of high vorticity magnitude,  $|\boldsymbol{\omega}| \geq 64.8$ , isolates over 120,000 distinct regions of strong vorticity (see Figure 2a). The global histogram of helicity (see Figure 3a) shows a nearly uniform distribution, indicating that all values of the angle between the velocity and vorticity vectors occur in the volume with similar frequency. A visualization dependent histogram, restricted to areas of high vorticity magnitude, shown in Figure 3b, yields similar insights with a nearly uniform helicity distribution. However, an examination of feature-local helicity distributions shows that in any one vortical structure (defined by this same vorticity threshold) the distributions do not exhibit this uniformity! In fact, we find that individual regions of high vorticity magnitude can be distinguished by unique helicity distributions signatures, as shown in Figure 1. Three distinct populations are apparent from the set of feature-local helicity histograms: wide noisy distributions (see row #7 in Figure 1), distributions trending toward high absolute values of helicity (see



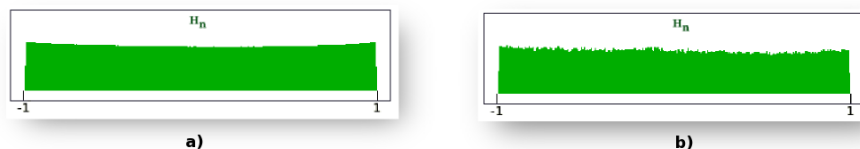
**Fig. 1.** Structure analysis showing histograms representing the local distributions of normalized helicity and velocity magnitude on each structure. The user can incorporate any data field or derived field into this analysis. In addition to the local distribution, the central moments of the distributions are calculated and can be used to sort, filter, select, and classify the individual structures. Multiple fields can be analyzed, displaying the relationships between the local distributions. Notice the helicity distributions of individual connected regions of high vorticity magnitude have very distinctive signatures. For example, compare rows #4 and #13. Both have relatively narrow distributions, but row #4 has a mean near zero and row #13 a mean near one, suggesting unique flow dynamics in the two regions.

rows #11 & #12 in Figure 1), and distributions that peak near low values of helicity (see row #4 in Figure 1).

By isolating structures using our transfer function, we were able to determine that the wide noisy distributions are of tangled structures (see Figure 2b) whose components were not well separated by the initial opacity threshold used in the selection step. Distributions that peak near minimum and maximum absolute helicity values correspond to crisp isolated tube-like structures. Those with low absolute helicity are dominated by flows with nearly orthogonal velocity and vorticity vectors. Those with high absolute helicity are dominated by flows with nearly parallel velocity and vorticity vectors, oriented either in the same (positive values) or opposite (negative values) directions. Volume renderings of the vorticity of the two populations are qualitatively indistinguishable – the features



**Fig. 2.** Progressive refinement through selective iteration: a) a transfer function is used to isolate regions of strong vorticity (shown with a visualization dependent histogram of helicity); b) one of the connected regions is selected, isolated, and visualized (shown with the feature-local histogram of helicity); c) the opacity threshold on the vorticity axis of the transfer function is increased and then sub-structures of the previously isolated “super-structure” are isolated (shown with feature-local histograms of helicity).

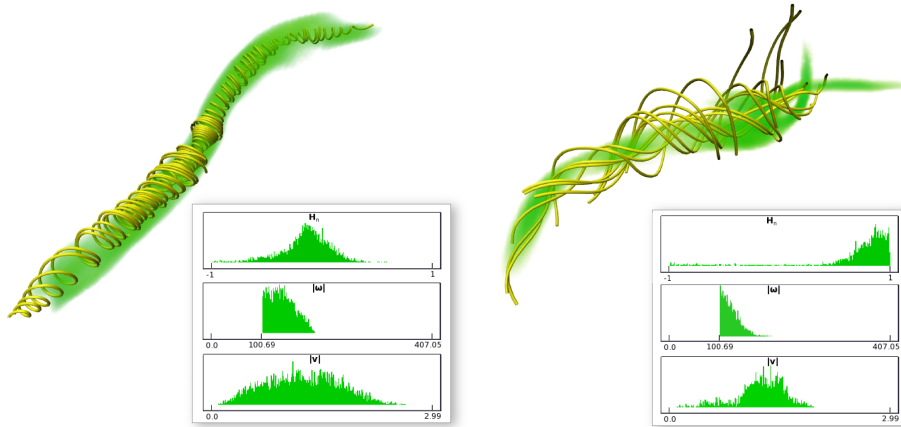


**Fig. 3.** Histograms of helicity shows a nearly uniform distribution of angles between the velocity and vorticity vectors. a) Helicity over the entire data set. b) Helicity values in regions of high vorticity.

are of similar sizes and shapes. However, streamlines seeded within these regions highlight their differences (see Figure 4). In the regions dominated by low helicity values, streamlines follow the vortex structure in a very tight, compact winding. These vortical structures are dominated by twist. By contrast, velocity streamlines seeded in the regions of predominately high helicity follow the vortex structure in loose open windings. These vortical structures are dominated by writhe. While both of these regions have high helicity values along the core, the twisting feature is dominated by low helicity values (see Gruchalla et al. [22] for further analysis).

By iterating our visualization and analysis pipeline, we can further deconstruct the complex features with the broad noisy helicity distributions into sub-structures and investigate their individual local distributions (see Figure 2). After we have identified a “super-structure” by its wide noisy helicity distribution, we select and visualize that structure in isolation. Then, by increasing the vorticity threshold, we begin to separate the individual sub-structures within the

isolated super-structure. A second iteration of the connected-component analysis, focused on this region, identifies these individual structures. Finally, by computing and visualizing the individual feature-local histograms, we again see that once the vorticity threshold is set to isolate tube-like regions, these regions have unique helicity distributions (see Figure 2c).

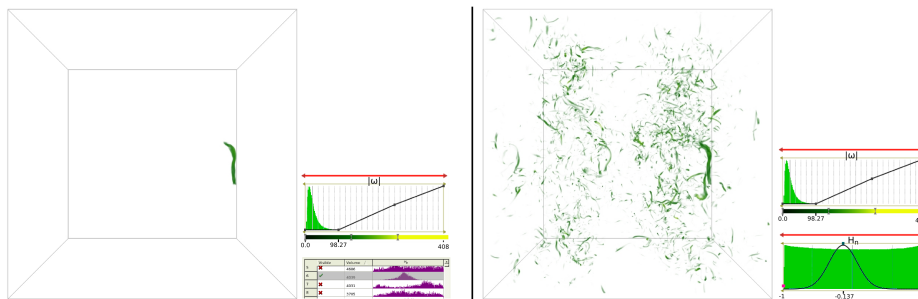


**Fig. 4.** Two volume and stream-line renderings accompanied by their feature-local histograms of helicity, vorticity magnitude, and velocity magnitude. Left: Streamlines in a feature isolated by low absolute mean helicity follow the structure in a tight winding. This type of structure is dominated by twist. Right: Streamlines seeded in a feature isolated by high absolute mean helicity follow the structure in a loose open helix. This type of structure is dominated by writhe.

#### 4.1 Result Comparison

A comparison of our results to a standard two-dimensional transfer function based on global distributions demonstrates the utility of incorporating feature-local information into the transfer function. Visualizing regions of high vorticity with high or low helicity using standard two-dimensional transfer function isolates a different set points than our transfer function and generates a significant amount of noise. Some individual structures, as is clear from Figure 1, have helicity distributions that are dominated by high or low values but have tails representing a full range of helicity values. A two-dimensional transfer function that renders regions of high vorticity and low helicity will render *all* the low helicity voxels, including those contained in structures dominated by high helicity values. For example, the twisting feature shown in Figure 4 was isolated with  $|\omega| \geq 98.27$  and has a local helicity distribution with a mean of  $-0.137$  and a standard deviation of  $0.281$ . We can attempt to isolate this feature using a

two-dimensional transfer function using the same opacity threshold on  $|\omega|$  and a Gaussian opacity function on helicity with  $\mu = -0.137, \sigma = 0.281$ , as shown in Figure 5. The result does visualize portions of the feature in question; however, it also visualizes portions of 7,395 other features that overlap those opacity settings! Those 7,395 visualized regions include regions of twisting and writhing, as well as regions that do not appear to constitute a vortex. This is because the two-dimensional transfer function is non-spatial: all areas of strong vorticity and low helicity are being visualized, without regard to local topology. Similarly, high helicity values contained in structures dominated by low helicity would not be rendered, fragmenting what our analysis finds to be complete and physically sensible structures. Our new transfer function framework provides a fundamentally different approach: both local and global information are incorporated into the visualization selection process, enriching the transfer function design process with spatial information.



**Fig. 5.** Comparison of a feature-local two-dimensional transfer function versus a standard two-dimensional transfer function. Left: a volume rendering of a twisting feature, isolated using a global vorticity magnitude transfer function coupled with a feature-local helicity transfer function. The feature-local helicity distribution has a mean of  $-0.137$  and a standard deviation of  $0.281$ . Right: A volume rendering using a standard two-dimensional (vorticity magnitude versus helicity) transfer function with a Gaussian opacity function in the helicity dimension with  $\mu = -0.137, \sigma = 0.281$ . The portions of the twisting feature shown on the right are visualized; however, so are the absolute low values of helicity from 7,395 regions, including regions near writhing vortices and regions without apparent vortices.

## 5 Conclusion

In this paper, we have described a powerful selective visualization system that provides the user the facilities to interactively investigate multivariate relationships *local* to individual features. Our system incorporates imaging and visualization techniques. We introduced a systematic integration of feature-local histograms and statistics into the analysis and visualization pipeline. We provide



the ability to progressively refine the features and in turn, provide a method to investigate multivariate relationships at multiple scales. This capability is particularly useful when attempting to segment ill-defined features, such as those in turbulent flow simulations where no formal definitions for features or structures of interest exist. We have demonstrated that a versatile, multivariate, feature-local approach is advantageous when investigating ill-defined features. The reasoning behind this is two-fold. First, feature-local multivariate distributions may be distinctive where global distributions are not, providing insight into the feature properties. Second, this type of approach allows for the comparative investigation of feature-identification techniques by examining the sensitivity of the structure properties to those techniques. We have demonstrated the power and the efficacy of our approach with an analysis of Taylor-Green turbulence, which resulted in the discovery of two populations of strong vorticity structures with distinct flow dynamics. These populations were discovered by a systematic examination of helicity histograms calculated local to structures defined as regions of high vorticity magnitude. These two feature populations were unlikely to be discovered by traditional visualization approaches that are based on histograms of global distributions, since helicity values are uniformly distributed across all values of vorticity magnitude in these data. In addition, our case study makes a strong case for visualization-driven analysis. These data are characterized by tens of thousands of vortical features, hindering traditional approaches to statistical overviews. Our system allowed interactive investigation of feature-local histograms on multiple data fields. When feature-local helicity histograms were visualized, unique populations of features were immediately apparent. Hypotheses on the nature of dynamics of these populations were then immediately and interactively tested by visualizing individual features and seeding streamlines within their boundaries.

## References

1. Mininni, P.D., Alexakis, A., Pouquet, A.: Nonlocal interactions in hydrodynamic turbulence at high reynolds numbers: the slow emergence of scaling laws. *Physical review. E, Statistical, nonlinear, and soft matter physics* **77** (2008)
2. Silver, D., Zabusky, N.J.: Quantifying visualizations for reduced modeling in nonlinear science: Extracting structures from data sets. *Journal of Visual Communication and Image Representation* **4** (1993) 46–61
3. Silver, D., Wang, X.: Tracking and visualizing turbulent 3d features. *IEEE Transactions on Visualization and Computer Graphics* **3** (1997) 129–141
4. Ebling, J., Scheuermann, G.: Clifford convolution and pattern matching on vector fields. In: *Proceedings of IEEE Visualization*. (2003) 193–200
5. Heiberg, E., Ebbers, T., Wigstrom, L., Karlsson, M.: Three-dimensional flow characterization using vector pattern matching. *IEEE Transactions on Visualization and Computer Graphics* **9** (2003) 313–319
6. Helman, J.L., Hesselink, L.: Representation and display of vector field topology in fluid flow data sets. *Computer* **22** (1989) 27–36
7. Theisel, H., Weinkauff, T., Hege, H.C., Seidel, H.P.: Saddle connectors - an approach to visualizing the topological skeleton of complex 3d vector fields. In: *Proceedings of IEEE Visualization, IEEE Computer Society* (2003) 225–232

8. Scheuermann, G., Tricoche, X.: Topological methods for flow visualization. In Hansen, C., Johnson, C., eds.: *Visualization Handbook*. Academic Press (2005)
9. Post, F.H., Vrolijk, B., Hauser, H., Laramée, R.S., Doleisch, H.: The state of the art in flow visualisation: Feature extraction and tracking. *Computer Graphics Forum* **22** (2003) 775–792
10. Suzuki, K., Horibata, I., Sugie, N.: Linear-time connected-component labeling based on sequential local operations. *Computer Vision and Image Understanding* **89** (2003) 1–23
11. Wong, P.C., Bergeron, R.D.: 30 years of multidimensional multivariate visualization, IEEE Computer Society (1997) 3–33
12. Bürger, R., Hauser, H.: Visualization of multi-variate scientific data. In: *Proceedings of EuroGraphics 2007 (State of the Art Reports)*. (2007) 117–134
13. Jänicke, H., Wiebel, A., Scheuermann, G., Kollmann, W.: Multifield visualization using local statistical complexity. *IEEE Transactions on Visualization and Computer Graphics* **13** (2007) 1384–1391
14. Sauber, N., Theisel, H., Seidel, H.P.: Multifield-graphs: An approach to visualizing correlations in multifield scalar data. *IEEE Transactions on Visualization and Computer Graphics* **12** (2006) 917–924
15. Kniss, J., Kindlmann, G., Hansen, C.: Multidimensional transfer functions for interactive volume rendering. *IEEE Transactions on Visualization and Computer Graphics* **8** (2002) 270–285
16. Park, S.W., Budge, B., Linsen, L., Hamann, B., Joy, K.I.: Multi-dimensional transfer functions for interactive 3d flow visualization. In: *Proceedings of the 12th Pacific Conference on (PG'04) Computer Graphics and Applications*, IEEE Computer Society (2004) 177–185
17. Bajaj, C.L., Pascucci, V., Schikore, D.R.: The contour spectrum. In: *Proceedings of the 8th conference on Visualization '97*, IEEE Computer Society Press (1997) 167–175
18. Tenginakai, S., Machiraju, R.: Statistical computation of salient iso-values. In: *Proceedings of the symposium on Data Visualisation 2002*, Barcelona, Spain, Eurographics Association (2002) 19–24
19. Correa, C., Ma, K.L.: Size-based transfer functions: A new volume exploration technique. *IEEE Transactions on Visualization and Computer Graphics* **14** (2008) 1380–1387
20. Lundström, C., Ljung, P., Ynnerman, A.: Local histograms for design of transfer functions in direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics* **12** (2006) 1570–1579
21. Clyne, J., Mininni, P.D., Norton, A., Rast, M.: Interactive desktop analysis of high resolution simulations: application to turbulent plume dynamics and current sheet formation. *New Journal of Physics* **9** (2007)
22. Gruchalla, K., Rast, M., Bradley, E., Clyne, J., Mininni, P.: Visualization-driven structural and statistical analysis of turbulent flows. In: *Advances in Intelligent Data Analysis VIII*. Volume 5772 of *Lecture Notes in Computer Science*. Springer (2009) 321–332
23. Kindlmann, G., Durkin, J.W.: Semi-automatic generation of transfer functions for direct volume rendering. In: *Proceedings of IEEE Visualization*. (1998) 79–86
24. Kniss, J., Premoze, S., Ikits, M., Lefohn, A., Hansen, C., Praun, E.: Gaussian transfer functions for multi-field volume visualization. In: *Proceedings of IEEE Visualization*, IEEE Computer Society (2003) 497–504