Towards New Grounds in Visualization

Helwig Hauser, Hauser@VRVis.at,
VRVis Research Center, http://www.VRVis.at/
March 2005

At the IEEE Visualization Conference 2004 in Austin, Texas, there was a lot of discussion about the future of visualization research, especially about the future of scientific visualization. The fact that two visualization communities (SciVis and InfoVis) develop almost parallel to each other (instead of interweaving a bit) also raised questions, especially on the SciVis side. In this article this discussion is resumed for a moment before the great potential of an integration of good solutions from both fields is demonstrated. A new technology called SimVis is briefly introduced for the interactive visual analysis of (scientific) simulation data, which is based on an intelligent combination of approaches both from InfoVis as well as also from SciVis. Since the integration of SciVis and InfoVis is only one interesting option for visualization research, also other potential directions are briefly addressed in this paper.

Bridging a Gap First

Scientific Visualization (SciVis) and *Information Visualization* (InfoVis) are – unfortunately – well-established terms in the visualization community. This is unfortunate, because, first of all, the terms are misleading themselves (InfoVis is also scientific and SciVis is also about visually communicating information to the user). Next, the two terms are much too often mistaken as delimiting distinct problem domains – instead, they rather delimit two visualization communities (sure, some of the problems which are dealt with in the two fields are different from each other, but in general the goals are very much the same). Finally, even the traditional definitions (SciVis deals with scientific data whereas data is abstract in the InfoVis domain) do not really help. There are cases in the InfoVis domain which deal with data with a strong reference to space and time (for example InfoVis of geospatial data [FS04, KW04]) as well as data exhibiting concrete physical form (e.g., music visualization [W02, PGW03] or InfoVis of biomolecular data [AJR03, SND04]). On the other hand, InfoVis techniques also are successfully applied to visualize scientific data [GR+00, HLD02, KBH04, PKH04, RE04, LT+05]. Accordingly, we should be careful when using the two terms to characterize what we are doing.

Further below in this paper the successful integration of InfoVis techniques in a SciVis application is demonstrated (section about SimVis), but beforehand a few comments should be made on other interesting future opportunities for visualization research (with a special focus on options for SciVis).

New Challenges in Visualization

In the context of Scientific Visualization, we have come to a point, where new challenges are to be addressed. In volume rendering, for example, very good solutions are already available – by now we pretty well know how to render three-dimensional data to flat screens, either in software or on the GPU, with a reasonable number of different projection techniques (from semi-transparent volume rendering to non-photorealistic volume illustration) as well as making use of different kinds of optimizations, e.g., for the rendering of large and time-dependent volume data. Due to the continuously increasing amounts of data to be visualized, the question of *what* to show (or what *not* to show) becomes increasingly important as compared to the question of *how* to visualize some data. In the future, we will also have to deal with a number of new challenges in this field, including

(a) the efficient and application-dependent extraction of features prior to visualization, e.g., to utilize object segmentation [HM+01] or multi-dimensional feature characterization [KKH01] in visualization – this addresses the question of what to show as well as the question of how to develop a semantic layer on top of the otherwise "pure" data –,

(b) the focus+context visualization of 3D (and also 4D = 3D + time) scientific data based on data semantics and a system-inherent representation of the user's interest – this addresses the integration of high-level access methods to data within the visualization so that the user can address the data in his/her own words –,

(c) the visualization of multi-variate (or multi-valued) datasets as, e.g., resulting from the registration of multiple datasets from different acquisition modalities or as resulting from a heterogeneous data repository,

(d) the research of usability issues in visualization, e.g., with respect to human perception and cognition [W04], knowledge crystallization [CMS99], and sensemaking [C04], or through the pursuit of user studies [KH+03].

In the following, the 3D focus+context visualization of multi-dimensional data from CFD simulation is demonstratead as one example for the successful integration of InfoVis and SciVis. It is illustrated how useful concepts from the InfoVis domain, such as focus+context visualization, linking&brushing, etc., can be adopted by SciVis research to improve results.

Interactive Visual Analysis with SimVis

There is a provocative claim to make: in many cases, scientific datasets are more rewarding to deal with in InfoVis than truly abstract data. Due to the usually continuous character of scientific data, e.g., the temperature distribution across a simulation domain or the distribution of data values in a CT dataset, many InfoVis-plots often are more efficient in terms of how well the visualization space is used for this kind of data than for data such as questionnaire evaluations, CRM data, etc. (the continuous character of scientific data very well matches the continuous characteristics of a screen – in the spatial, the temporal, and also in the chromatic direction). In a system called SimVis a number of technological approaches from the InfoVis domain have been exploited to improve the visual analysis of scientific data:

- Selected InfoVis views such as histograms [KBH04], scatterplots [PKH04], and parallel coordinates [HLD02], are used to put whatsoever data attributes in a visual relation to each other – the user can use (example 1) a histogram to investigate the distribution of velocity values in a CFD simulation, or (ex. 2) use a 3D scatterplot to visually analyze the interrelation of the dominant flow direction, i.e., a spatial dimension, and two selected flow attributes (for example a measure for turbulence in the flow and flow velocities), or (ex. 3) use a plot of parallel coordinates to investigate the multi-dimensional relations between several flow attributes.

  In Fig. 1[1], a CFD dataset is visualized by the means of a histogram (d, wrt. velocity values), a 2D scatterplot (b, velocities on x, turbulence values on y), a 3D scatterplot (c, velocities, turbulence and temperature values), and a plot of parallel coordinates (a, eight different flow attributes shown). In the two scatterplots and in the histogram brushes are used to mark subsets of special interest. In Fig. 1/a, one axis (showing velocity values) is used to color the polylines – this eases to relate data items across multiple dimensions.

- An approach called *linking & brushing* in the InfoVis domain [BM+91, W94] is used in SimVis to allow for interactive visual analysis of high-dimensional and complex datasets [DGH03, DM+04, DM+05]. Multiple views, often also of different kind, are used to all jointly display one dataset under investigation [BWK00]. Brushing enables the user to interactively mark certain subsets of the data as being of special interest – the user directly brushes the displayed data in one of the views [BC87, MW95, W96, C03] or uses off-

[1] All related figures are collected at the end of this paper.

screen sliders and direct numerical input widgets [S94] to tell the system what he or she (currently) focuses on. Linking between views ensures that the entire visualization is visually consistent, i.e., in all the views the same data subsets are consistently visually enhanced (or deemphasized), for example, by everywhere coloring the data in focus red.

In Fig. 2 three scatterplots and three histograms have been used to specify the current focus in the visualization of a simulation of hurricane Isabel (cloud structures and land). The scatterplots (ca), (cv), and (av), and the histogram (h) are used to highlight all those parts of the clouds in the hurricane which exhibit a significant amount of wind speed (cv), which do not raise up too high into the atmosphere (h), and which either move noth or south (two 3D brushes in ca and av). The distinction between north and south winds enables to detect an interesting flow behaviour in the north of the hurricane. For the purpose of additional orientation, a (dimmed) visualization of the east cost of America has been added through (partly) selecting the respective cells in two histograms (h' & h''). The 3D view (3D) provides a spatial and also time-dependent 3D visualization (color visualizes wind directions, green: north, purple: south).

- The interactive process of semantically annotating the data (through *brushing*) to differentiate between more and less interesting parts of the data consequently causes the views to show a (so called) *focus+context visualization* of the data [H05]. Data subsets *in focus* are visually enhanced as compared to their *context*, i.e., the rest of the data, which is just shown in reduced form, e.g., less opaque and not colored, to provide a visual reference for the current focus and to ease orientation and navigation. In SimVis, the approach of focus+context visualization also has been extended to the more traditional SciVis views, i.e., the 3D rendering of the data. Data in focus is visually discriminated from the context by selective coloring as well as by variations of opacity and style.

  Fig. 3 demonstrates how focussing allows to reveal interesting structures in a 3D dataset. In (F+C) the interface between water and air has been selected in a two-phase simulation of the flood after the breaking of a dam (on the right) and its effect on an obstacle in the mid-right of the flow domain (color: flow velocities). In (C) no features have been selected whereas in (F) all the cells are shown in the same style – whereas in (C) too little is shown, in (F) too much is visualized.

  In Fig. 4, all those cells in a diesel particulate filter (DPF) for passenger cars have been highlighted which are characterized by the presence of lots of CO and CO_2 and which also exhibit relatively high temperatures (color shows flow velocities) – thereby the front of an oxidation process is shown in the DPF which is used for the periodic regeneration of the filter [DM+04].

  In Fig. 5, all those cells in a 45° subset of the combustion chamber (in one cylinder of a diesel engine) are highlighted (at four different time steps of the simulation) where the combustion already is well progressed while still a significant amount of diesel is left (an undesired situation) – color visualizes the amount of O_2 in the cells. The respective analysis [DM+05] affirms that combustion stops too early because too little oxygen is collocated with the fuel mixture to be burnt in certain locations.

- In addition to the above mentioned points (using InfoVis techniques to visualize attribute space of scientific data, establishing a linking&brushing framework for interactive data analysis, and using focus+context visualization to incorporate a notion of user interest in the visualization), the importance of interaction should be pointed out again. Ben Shneiderman's visual information seeking mantra [S96] well describes the main character of working with SimVis: First, usually sorts of overview visualization are important – the user wants to get oriented with respect to the data, understand the main features, etc. Then, as soon as specific questions about the data arise, the user wants to perform an information drill-down which allows to gradually refine a feature specification (both graphically on screen as well as numerically), before at a certain point of the detailed analysis the user might want to export the results of the analysis, for example, to compare a result with another analysis of another dataset.

The above mentioned examples of using technological results from one field to improve results in another most probably have model character for other cases. Since the prime focus of this paper is on how to conquer new grounds in visualization, the discussion of SimVis is limited to the above mentioned points here, even though there are other technological aspects of interactive visual data analysis with SimVis, e.g., the smooth delimitation of features from the respective context or the use of a feature definition language to explicitly represent an analysis session, which also are responsible for the success of this technological approach (winner of the IEEE Visualization Contest 2004, best paper at SimVis 2005, many successful case studies, etc.).

Conclusions

At the end of this article, a couple of personal thoughts are recalled which potentially are important for future visualization research and a few conclusions are drawn from a personal point of view (of course this is a controversial compilation, but that is what it also should be). Visualization (and especially Scientific Visualization) seems to be at a point where new challenges are to be addressed. It seems that we should increasingly care about the question of what to show (instead of how to show it). Accordingly, we probably will have to start doing more about data semantics – additionally with then also considering such high-level aspects of data during visualization – one potentially fruitful future direction is the development of visualization interfaces which enable the user to address his/her data in his/her own (application) terms. The SimVis approach, as briefly introduced above, can act as one example for aiming at this goal here. Furthermore, it also seems that we should care more about the actual abilities of human perception and cognition – there is a lot of research about what human observers can actually perceive best [W04]. And we will most probably also have to do more about evaluation, e.g., by the means of user studies [KH+03]. Finally, this wrap-up is (again) focused on sensemaking. In most cases visualization only is part of a bigger application system. Accordingly, it usually has to fulfill a very specific task (or more of them) in the context of some larger application. So it is very useful to regularly recall this larger embedding of visualization, especially when doing visualization research (rendering beautiful images is nice but usually not the first-hand task).

One conclusion to draw is that to successfully engage with most of these new challenges, visualization researchers will have to become more interested in neighboring, related fields of research – for (a) feature extraction, for example, computer vision already provides a set of very interesting approaches; for (b/c) focus+context visualization and the visualization of multi-dimensional data, information visualization offers promising supply; and for (d) user-centered visualization, the fields of usability, HCI (human-computer interaction), perception, and cognition research do offer a lot of useful results.

Acknowledgements

This paper is about work which (in parts) has been done in the basic research on visualization in the VRVis Research Center in Vienna, Austria, which is funded by an Austrian governmental research program called Kplus. The author thanks the following colleagues (in alphabetical order) who all contributed to this research: Helmut Doleisch, Martin Gasser, Robert Kosara, Michael Mayer, Matej Mlejnek, Philipp Muigg, Harald Piringer, and several others.

References

[AJR03] Alston, Mark; Johnson, Colin; Robinson, Gary: Colour Merging for the Visualization of Biomolecular Sequence Data. Proceedings of the 7th International Conference on Information Visualization (IV '03), pp. 169–175, 2003.

[BWK00] Baldonado, Michelle; Woodruff, Allison; Kuchinsky, Allan: Guidelines for Using Multiple Views in Information Visualization. Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '00), pages 110–119, 2000.

[BC87] Becker, Richard; Cleveland, William: Brushing Scatterplots. Technometrics, 29(2):127–142, 1987.

[BM+91] Buja, Andreas; McDonald, John; Michalak, John; Stuetzle, Werner: Interactive Data Visualization Using Focusing and Linking. Proceedings of IEEE Visualization (Vis '91), pp. 156–163, 1991.

[CMS99] Card, Stuart; Mackinlay, Jock; Schneiderman, Ben: Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann Publishers, 1999.

[C04] Card, Stuart: From Information Visualization to Sensemaking: Connecting the Mind's Eye to the Mind's Muscle. Keynote Talk & Proceedings of the 10th IEEE Symposium on Information Visualization (InfoVis '04), p. xii, 2004.

[C03] Chen, Hong: Compound Brushing. Proceedings of the IEEE Symposium on Information Visualization (InfoVis '03), pp. 181–188, 2003.

[DGH03] Doleisch, Helmut; Gasser, Martin; Hauser, Helwig: Interactive Feature Specification for Focus+Context Visualization of Complex Simulation Data. Proceedings of the 5th Joint IEEE TCVG – EUROGRAPHICS Symposium on Visualization (VisSym '03), pp. 239–248, 2003.

[DM+04] Doleisch, Helmut; Mayer, Michael; Gasser, Martin; Wanker, Roland; Hauser, Helwig: Case Study: Visual Analysis of Complex, Time-Dependent Simulation Results of a Diesel Exhaust System. Proceedings of the 6th Joint IEEE TCVG – EUROGRAPHICS Symposium on Visualization (VisSym '04), pp. 91-96, 2004.

[DM+05] Doleisch, Helmut; Mayer, Michael; Gasser, Martin; Priesching, Peter; Hauser, Helwig: Interactive Feature Specification for Simulation Data on Time-Varying Grids. Proceedings of Simulation and Visualization (SimVis '05), pp. 291–304, 2005.

[FS04] Fuchs, Georg; Schumann, Heidrun: Visualizing Abstract Data on Maps. Proceedings of the 8th International Conference on Information Visualization (IV '04), pp. 139–144, 2004.

[GR+00] Gresh, Donna; Rogowitz, Bernice; Winslow, Raimond; Scollan, David; Yung, Christina: WEAVE: A System for Visually Linking 3-D and Statistical Visualizations, Applied to Cardiac Simulation and Measurement Data.. Proceedings of IEEE Visualization (Vis '00), pp. 489–492, 2000.

[HM+01] Hauser, Helwig; Mroz, Lukas; Bischi, Gian-Italo; Gröller, Eduard: Two-Level Volume Rendering. IEEE Transactions on Computer Graphics and Visualization 7(3), pp. 242-252, 2001.

[HLD02] Hauser, Helwig; Ledermann, Florian; Doleisch, Helmut: Angular Brushing for Extended Parallel Coordinates. Proceedings of the 8th IEEE Symposium on Information Visualization (InfoVis '02), pp. 127–130, 2002.

[H05] Hauser, Helwig. Generalizing Focus+Context Visualization: Proceedings of the 2003 Dagstuhl Seminar on Scientific Computing, to appear, 2005.

[KW04] Kapler, Thomas; Wright, William: GeoTime Information Visualization. Proceedings of the 10th IEEE Symposium on Information Visualization (InfoVis '04), pp. 25–32, 2004.

[KKH01] Kniss, Joe; Kindlmann, Gordon; Hansen, Charles: Interactive Volume Rendering Using Multi-Dimensional Transfer Functions and Direct Manipulation Widgets. Proceedings of IEEE Visualization (Vis '01), pp. 255–262, 2001.

[KH+03] Kosara, Robert; Healey, Christopher; Interrante, Victoria; Laidlaw, David; Ware, Colin: Thoughts on User Studies: Why, How, and When. IEEE Computer Graphics & Applications 23(4):20–25, 2003.

[KBH04] Kosara, Robert; Bendix, Fabian; Hauser, Helwig: TimeHistograms for Large, Time-Dependent Data. Proceedings of the 6th Joint IEEE TCVG – EUROGRAPHICS Symposium on Visualization (VisSym '04), pp. 45–54, 2004.

[LT+05] Lowe, Nick; Tory, Melanie; Potts, Simeon; Datta, Amitava; Möller, Torsten: A Parallel Coordinates Style Interface for Exploratory Volume Visualization. IEEE Transactions on Visualization and Computer Graphics 11(1):71–80, 2005.

[MW95] Martin, Allen; Ward, Matthew: High-Dimensional Brushing for Interactive Exploration of Multivariate Data. Proceedings of IEEE Visualization (Vis '95), pp. 217–278, 1995.

[PGW03] Pampalk, Elias; Goebl, Werner; Widmer, Gerhard: Visualizing Changes in the Structure of Data for Exploratory Feature Selection. Proceedings of 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03), pp. 157–166, 2003.

[PKH04] Piringer, Harald; Kosara, Robert; Hauser, Helwig: Interactive Focus+Context Visualization with Linked 2D/3D Scatterplots. Proceedings of the 2nd International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV '04), pp. 49–60, 2004.

[RE04] Reina, Guido; Ertl, Thomas: Volume Visualization and Visual Queries for Large High-Dimensional Datasets. Proceedings of the 6th Joint IEEE TCVG – EUROGRAPHICS Symposium on Visualization (VisSym '04), pp. 255–260, 2004.

[SND04] Saraiya, Purvi; North, Chris; Duca, Karen: An Evaluation of Microarray Visualization Tools for Biological Insight. Proceedings of the 10th IEEE Symposium on Information Visualization (InfoVis '04), pp. 1–8, 2004.

[S94] Shneiderman, Ben: Dynamic Queries for Visual Information Seeking. IEEE Software 11(6):70–77, 1994.

[S96] Shneiderman, Ben: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Proceedings of the IEEE Symposium on Visual Languages (VL '96), pp. 336–343, 1996.

[W94] Ward, Matthew: XmdvTool: Integrating Multiple Methods for Visualizing Multivariate Data. Proceedings of IEEE Visualization (Vis '94), pp. 326–336, 1994.

[W04] Ware, Colin: Information Visualization: Perception for Design. Morgan Kaufmann Publishers, 2004.

[W02] Wattenberg, Martin: Arc Diagrams: Visualizing Structure in Strings. Proceedings of the 8th IEEE Symposium on Information Visualization (InfoVis '02), pp. 110–116, 2002.

[W96] Wills, Graham: Selection: 524,288 Ways to Say "This is Interesting". Proceedings of the 2nd IEEE Symposium on Information Visualization (InfoVis '96), pp. 54–61, 1996.

Figures

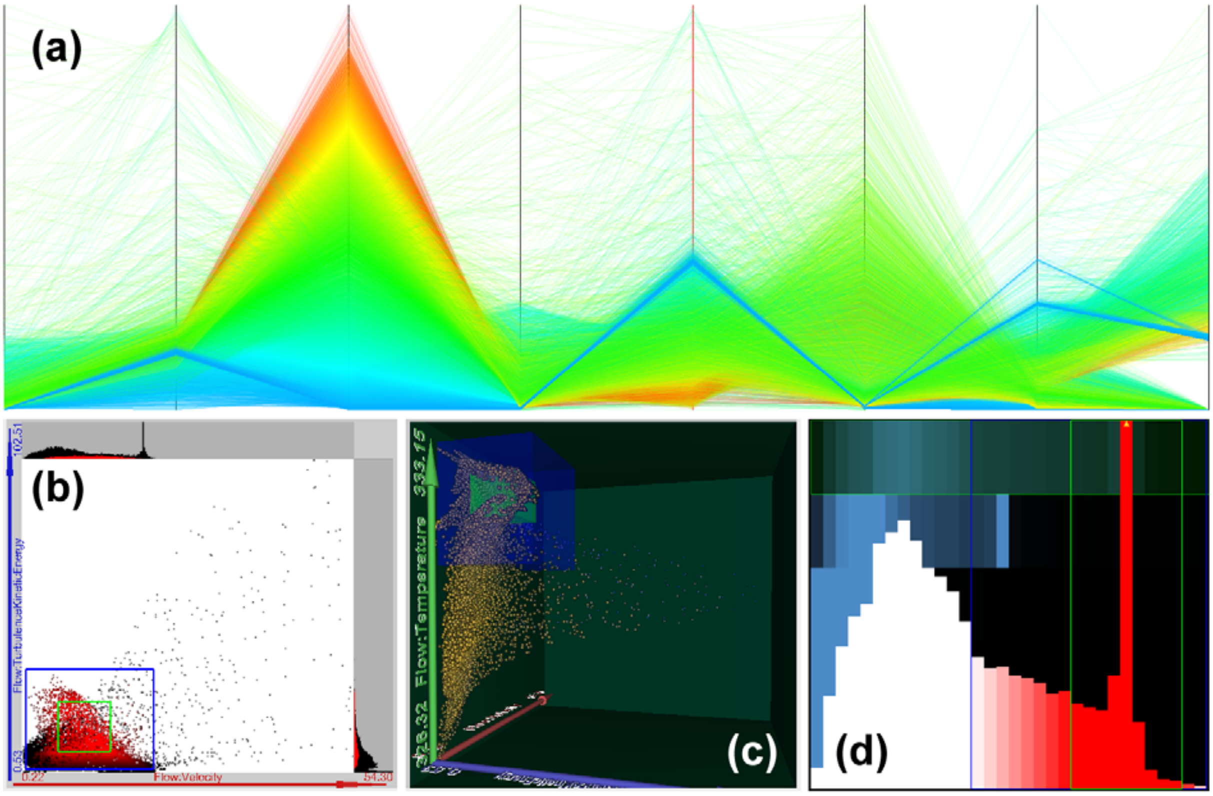

Figure 1: A CFD dataset is visualized by the means of a histogram of velocity values (d), a scatterplot "velocities vs. tubolence values" (b), a 3D scatterplot "velocities, tubolence, temperatures" (c), and parallel coordinates (a). Different aspects of the data become apparent through the use of different visualization techniques.

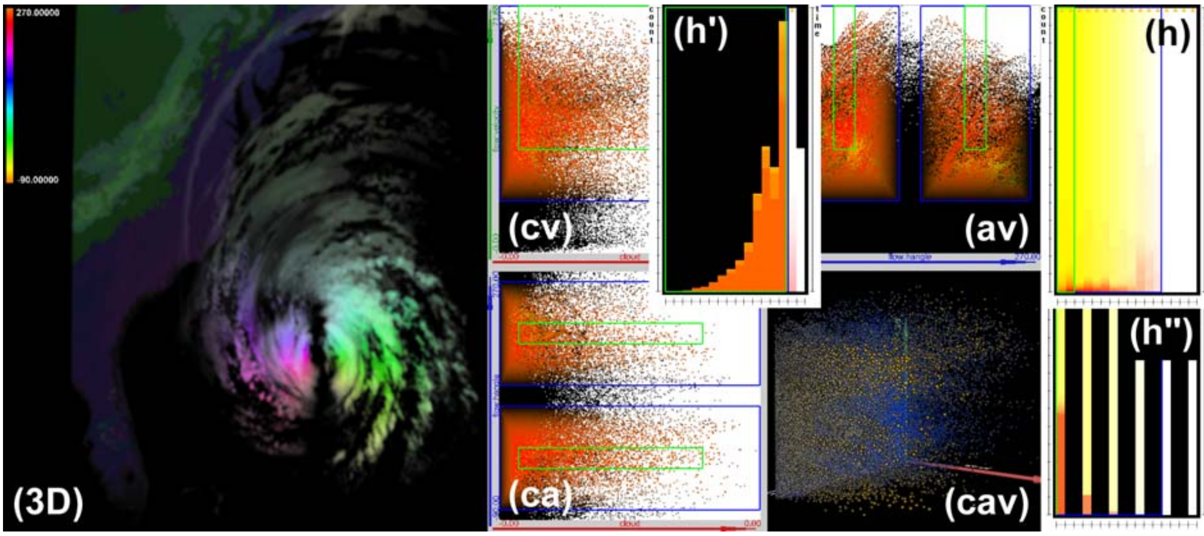

Figure 2: The simulated hurricane Isabel is visualized. Through the use of three scatterplots and a histogram, relatively fast clouds are emphasized which to do raise too high up into the atmosphere – only winds which are (at least to some extent) aligned north-south are highlighted to reveal an interesting flow structure in the north of the hurricane. Additionally the land is shown for reference.

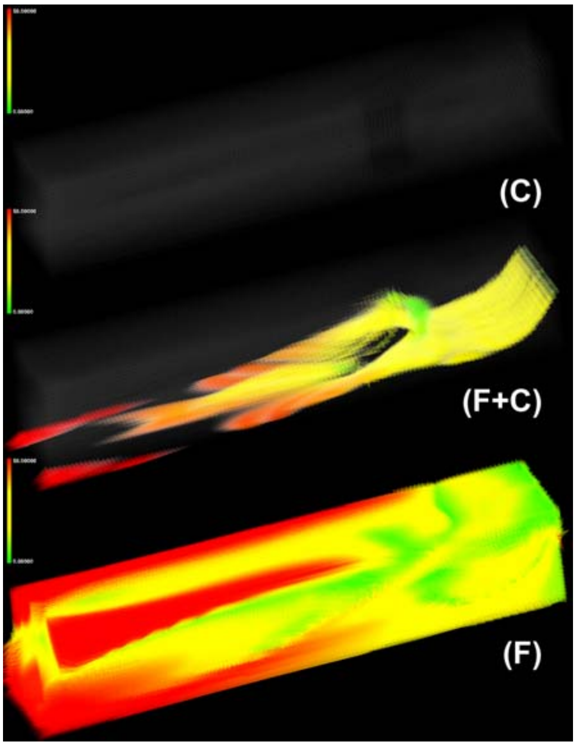

Figure 3: Example for the visual discrimination of data subsets in focus from the respective context. The flood as resulting from a breaking dam is shown together with its effect on an obstacle (in the mid-right of the flow domain). Only through the (interactive) focussing on specific subsets of the data (in this example the interface between water and air in “F+C”), meaningful images can be produced.

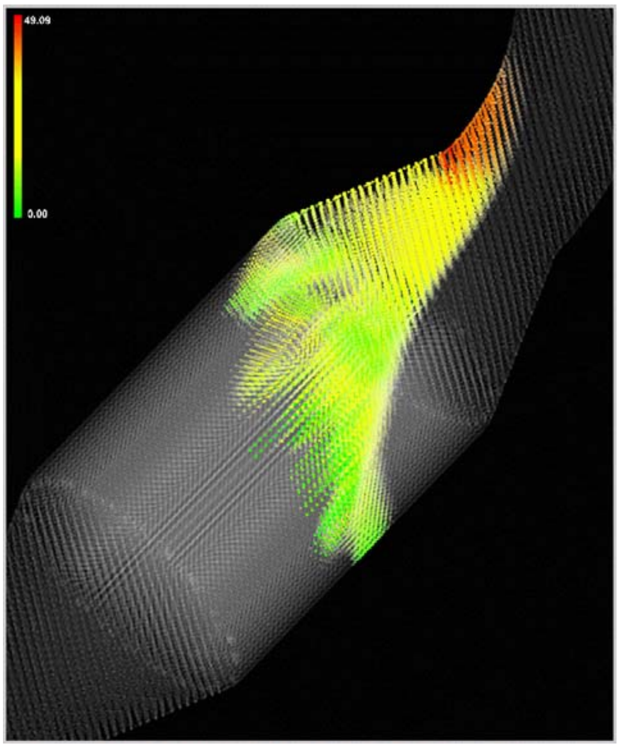

Figure 4: The front of an oxidation process in a diesel particulate filter (DPF) is shown [DM+04], which is used for the periodic regeneration of the filter – all those cells are highlighted which exhibit large amounts of CO and CO_2 while at the same time being relatively hot. Color shows flow velocity.

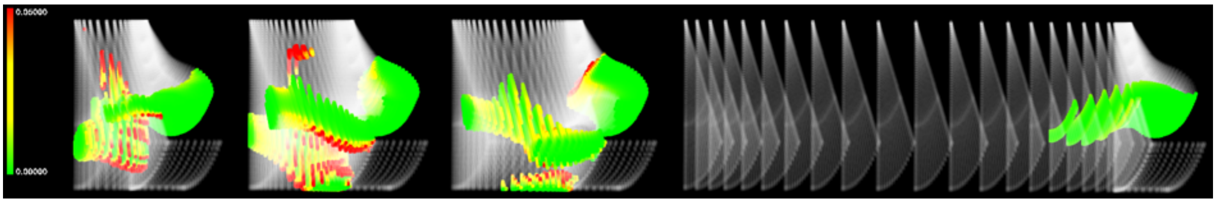

Figure 5: Four time steps of a combustion process are shown [DM+05], focussing on cells where the combustion is well progressed but where still a significant amount of diesel is left (undesired) – color shows the amount of oxygen at this cells (in cells with too little oxygen the combustions terminates too early).