Volume Data Mining Using 3D Field Topology Analysis

Issei Fujishiro and Yuriko Takeshima
Ochanomizu University and Research Organization for Information Science & Technology

Taeko Azuma
Ochanomizu University

Shigeo Takahashi
Gunma University

We took advantage of a 3D field topology analysis to automate visualization design for volume data mining. A feasibility study used a large-scale 4D simulated data set.

Volume visualization has served as an indispensable tool for exploring the inner structures and complex behavior of volumetric objects embedded in large-scale sampled or simulated 3D data sets. However, the rapid increase in data set sizes makes it difficult to adjust visualization-related parameters sufficiently for generating informative images. These representative parameters include the target level for isosurfacing and transfer functions for direct volume rendering. To compensate for the lack of interactivity—and to provide the user with the serendipity[1]—requires developing a mechanism for visual data mining that chooses appropriate values for the visualization parameters based on some available quantitative properties of a given volume data set.

This article proposes a novel approach to automating the settings of visualization parameter values for volume data mining. To this end, we extended the conventional Reeb graph-based approach to topological modeling of 3D surfaces[2,3] to capture the topological skeleton of a volumetric field. The analyzed results take the form of hyper Reeb graphs,[4] which give the basic reference structure for designing comprehensible volume visualization.

Representing volume field topology

To represent volume field topology, we can use an extension of Reeb graphs to 3D volume fields, hyper Reeb graphs. Let's look at the original concept first.

Reeb graphs

Shinagawa et al. originally imported the concept of the Reeb graph into computer graphics fields in order to reconstruct a topologically correct surface from cross-sectional contours extracted from computed tomography images.[2] They also applied the Reeb graph to represent the topological skeleton in their surface-coding system.[5] Others have used the Reeb graph and its variations for characterizing geographical features,[6,3] designing surfaces,[7] and extracting iso-contours/surfaces.[8]

The definition of the Reeb graph follows: Let h be the height function of a surface, and let p and q be points on the surface. The Reeb graph of h is obtained by identifying p and q, if the two points are contained in the same connected component on the cross section of the surface at the height $h(p)$ $(= h(q))$.

The nodes of the Reeb graph represent one of three kinds of critical points: a peak, a pass, or a pit. Let's consider the topological changes of equi-height contours at a peak, a pass, and a pit, respectively, from the uppermost contour to the lowermost. At a peak, a new contour appears, while an existing contour disappears at a pit. At a pass, a contour splits or two contours merge. In addition, the edge of the Reeb graph corresponds to a set of topologically equivalent connected contours on consecutive cross sections.

Figure 1 shows two simple surfaces, an ellipsoid and a torus, and their corresponding Reeb graphs. Note that for the torus, changing the directions of the height axis may give different Reeb graphs. This means that the Reeb graph provides an unambiguous, but nonunique, representation of the surface topology.

Hyper Reeb graphs

The hyper Reeb graph (HRG) extends the Reeb graph concept to 3D volume fields.[4] Theoretically, a volume decomposes into an infinite number of isosurfaces with different target values. We can capture the topological features of each isosurface by using the Reeb graph with a common direction for the height function. Therefore, by examining the sequence of isosurfaces in terms of the structure of Reeb graphs, we can find a particular field value—the *critical field value* (CFV)—for which the topological equivalence of consecutive isosurfaces is not maintained.

A hierarchical graph, HRG consists of two layers of topological data specification. The top layer of the graph is a linearly directed graph connecting, in ascending order, m nodes v_i with CFVs $f_i (i = 1, \ldots, m)$ and two boundary nodes v_0 and v_{m+1} with minimum and maximum field values f_0 and f_{m+1}, respectively. Each edge $e(v_i, v_{i+1})$ retains as its weight the corresponding topologically equivalent Reeb graph at the bottom layer as well as the length of the field interval $l_{i,i+1}$ $(= f_{i+1} - f_i)$. Note that if the field interval is open (closed), its boundary nodes are depicted with an open (solid) circle.

As a running example, let us consider the following analytical volume:

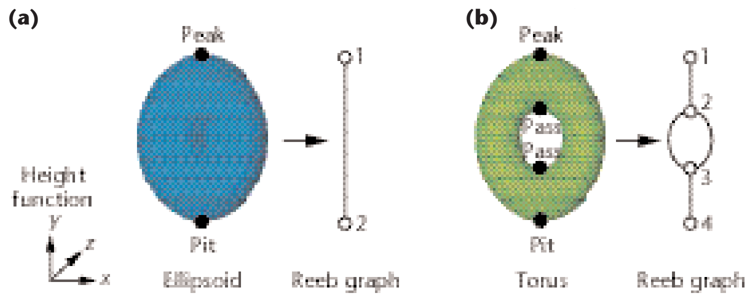
$$f(x, y, z) = a^2 + b^2$$

$$a = \exp\left(-\sqrt{x^2 + y^2 + z^2 + 1 - 2\sqrt{x^2 + y^2}}\right)$$

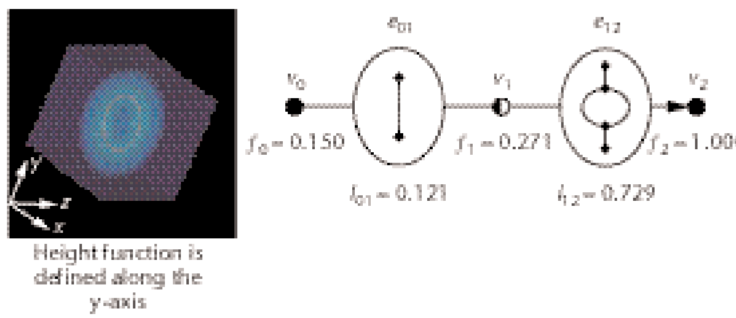
$$b = \exp\left(-\sqrt{x^2 + y^2 + z^2 + 1 + 2\sqrt{x^2 + y^2}}\right)$$

A simple arithmetic operation shows that the volume, called the *metatorus* hereafter, has a single CFV f_1 $(= 2e^{-2})$ when and only when $a = b$. Figure 2 illustrates how we construct an HRG for the volume. From the resulting HRG we observe that

- As the field value increases, ellipsoidal isosurfaces deflate, approaching a donut shape in the direction of the x-axis. Then, passing by the CFV, a single hole appears around the origin, to generate the sequence of nested tori.
- As the field value increases further, the diameter of the rounded tube of the tori becomes smaller.



1 **Reeb graphs for simple 3D surfaces: (a) ellipsoid and (b) torus.**



2 **Construction of an HRG for the metatorus volume.**

Comprehensible volume visualization

Now we attempt to take advantage of the HRG-based volumetric-field-topology description in order to enhance conventional volume visualization techniques.

Geometric object fitting

We consider the following two options:

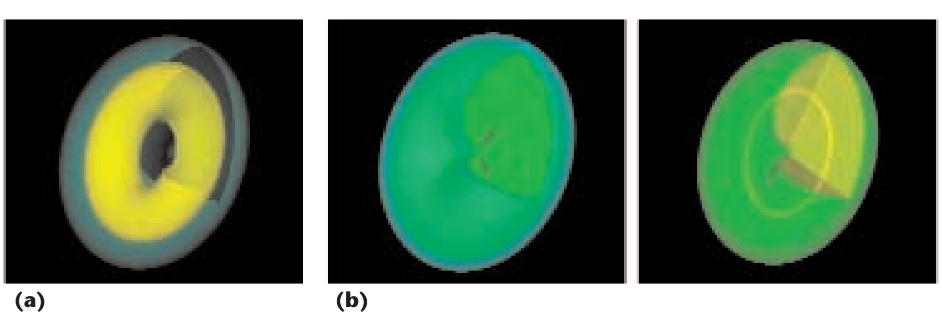
Method 1: Simultaneous display of $m+1$ semitransparent isosurfaces, each extracted with a field value $(f_i + f_{i+1})/2$ at the midpoint of the topologically equivalent field interval $[f_i, f_{i+1}]$ $(i = 0, \ldots, m)$. We can determine a plausible value for the opacity of each isosurface so as to reflect the mutual relationships among $l_{i,i+1}$ in order to understand the relative thickness of topologically equivalent field intervals.

Method 2: Decomposition of a given volume V into a sequence of $m+1$ nonoverlapping interval volumes $IV(f_i + f_{i+1})$ $(i = 0, \ldots, m)$; that is,

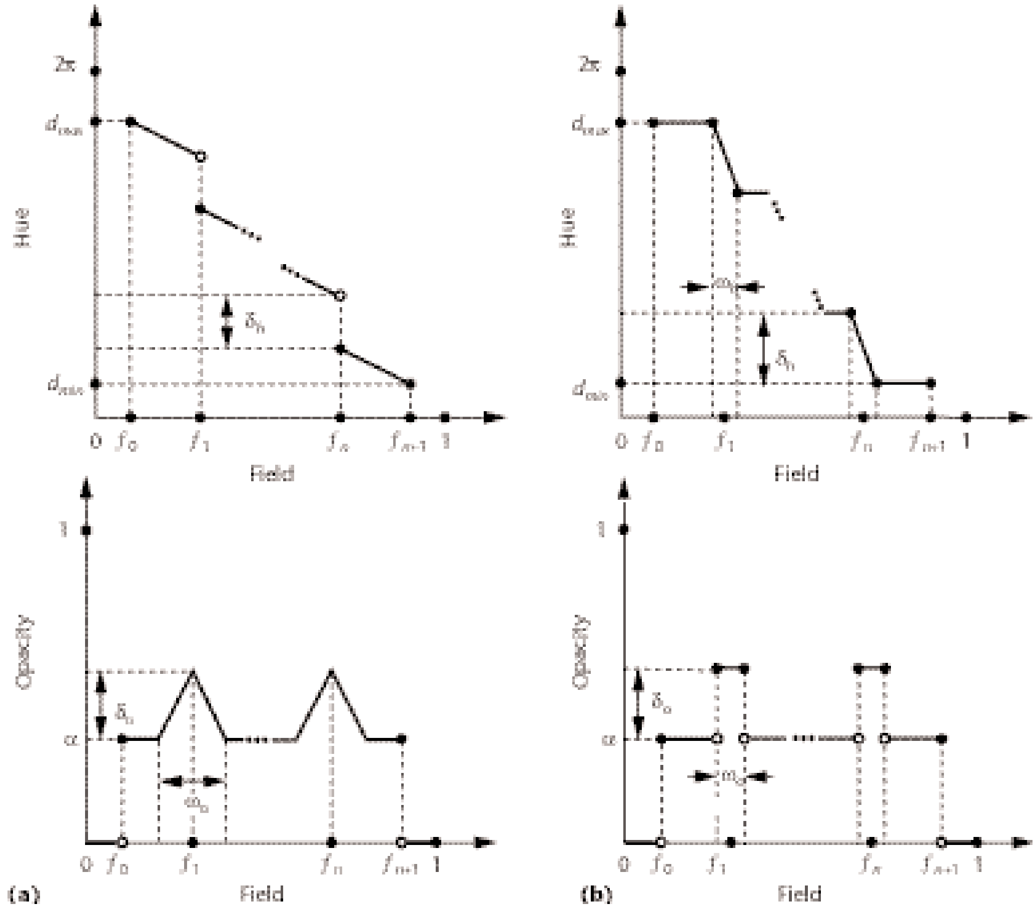
$$V = \bigcup_{i=0}^{m} IV(f_i, f_{i+1})$$

The *interval volume* was proposed as a solid data representation of a 3D subvolume for which the associated field values lie within a specified closed field interval.[9] Topological equivalence gives the rigid basis for the volume decomposition. In addition, the boundaries of each interval volume convey informative shapes of isosurfaces with CFVs, at exactly the location where the topology of level surfaces changes.

Figure 3 visualizes the metatorus volume with the above two methods in a comprehensible manner. The selected isosurfaces in Figure 3a can also serve as an effective set of basic frames for the flipbook approach to volume rendering. On the other hand, the set of interval volumes in Figure 3b should provide a good initial step for more sophisticated volume segmentation.



3 **Geometric object extraction from the metatorus volume based on the HRG in Figure 2. The upper-right-front octant of the volume is cropped. (a) Simultaneous display of two isosurfaces with 0.21 (ellipsoid) and 0.64 (torus). (b) Decomposition into two interval volumes, *IV*[0.15, 0.271] and *IV*[0.271, 1.0].**



4 Design principles of transfer function accentuation. (a) Principle 1 and (b) Principle 2.

Transfer function design

One of the most significant factors for determining the quality of volume-rendered images is the transfer function, which maps physical fields of a given volume data set to optical properties, such as color and opacity. Several methods of (semi-)automating transfer function design for informative volume rendering appear in the literature. These divide into two major categories, image-guided[10,11] or input-volume content-based.[12-14] The approach presented here, although novel, fits into the second category.

The basic idea of designing appropriate transfer functions based on the HRG is to accentuate the topological change in volume fields around CFVs in terms of both color (hue) and opacity. We will specify the two transfer functions within the analyzed subdomain $[f_0, f_{m+1}]$. Hue and opacity transfer functions are respectively set to be *undefined* and 0 (fully transparent) outside of the subdomain.

We consider the following two design principles:

Principle 1: The color transfer function is designed so that the rate of change in hue is uniform for all field values, except for a constant jump δ_h at each CFV f_i $(i = 1, \ldots, m)$. On the other hand, the opacity transfer function is designed to be a constant α (> 0), except for a common hat-like, small elevation around each CFV f_i $(i = 1, \ldots, m)$, whose height and width are δ_o and ω_o, respectively (Figure 4a).

Principle 2: The hue transfer function is designed to be elevated stepwise by a fixed amount δ_h, except for a linear change within a small interval of length ω_h around CFV $f_i (i = 1, \ldots, m)$. On the other hand, the opacity transfer function is designed to have a small fixed elevation δ_h relative to the base height α within the same interval of length ω_o $(= \omega_h)$ around CFV f_i $(i = 1, \ldots, m)$ (Figure 4b).

Figure 5 compares volume-rendered images of the metatorus volume with the two different designs of transfer functions. The transfer functions designed according to Principle 1 help in visually grasping topologically equivalent interval volumes. On the other hand, the transfer functions designed according to Principle 2 prove beneficial for observing the change in topological structures near CFVs in detail. In addition, this principle suits the bisection search for CFVs in digital settings because, in general, a CFV will likely be specified as an internal point belonging to a field interval of the minimum length.

Implementation

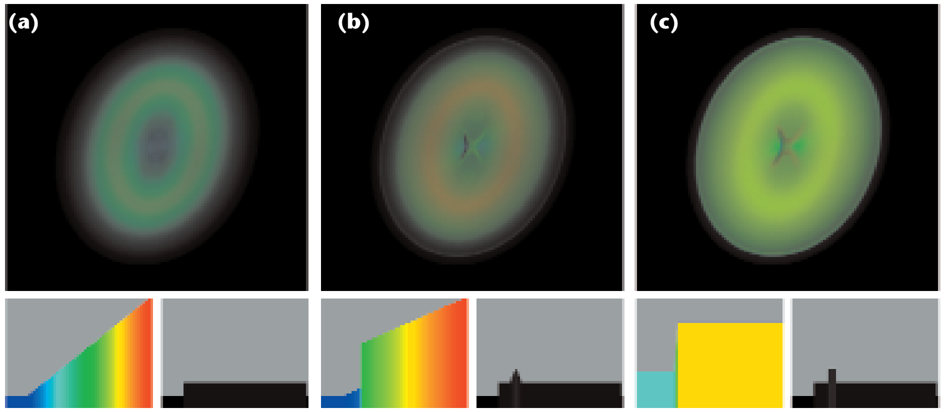
To make the present volume data mining methodology applicable to practical sampled or simulated data sets, we have developed a pilot environment on Advanced Visual Systems' visualization software platform AVS/Express version 5.0[15] running on an SGI O2 system (R5000 CPU, 180-MHz clock, and 192 Mbytes of RAM). All the experiments described here took place in the same environment.

Figure 6 shows a typical AVS/Express module network used for volume data mining. The HRG constructor is our main in-house module, which inputs a scalar volume and the axis direction for the height function. This module generates the corresponding HRG. Isosurfacing relies on an extended version of the Marching Cubes algorithm combined with an auxiliary algorithm, called Asymptotic Decider, for maintaining the topological consistency in triangle-patch connections.[9]

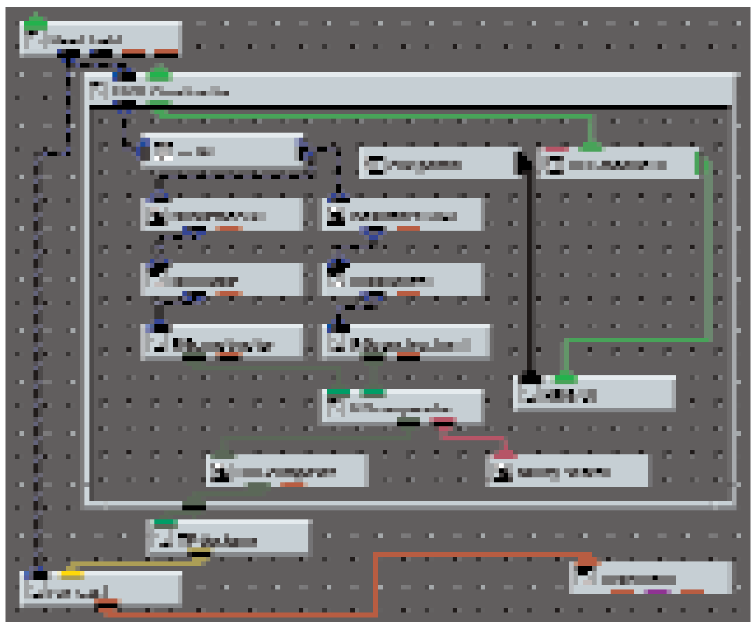
To reduce the large number of critical points—caused by high-frequency components likely to appear on an isosurface extracted from actual volume data sets—we adopt a Gaussian-type surface low-pass filter algorithm.[16] A robust algorithm[3] then constructs the Reeb graph of each isosurface, first extracting all the critical points on a surface correctly in the sense of Euler's formula. It then constructs the surface network[17] by tracing ridge and ravine lines, and finally converts the surface network to the corresponding Reeb graph. While the algorithm was originally limited to topological spheres,[13] we extended it so that it can characterize surfaces of arbitrary topological type.

The HRG constructor module employs a bisection-based method to find all the CFVs in the user-specified field interval $[f_0, f_{m+1}]$. To judge the homogeneity of Reeb graphs, we compute and compare two characteristic quantities of the graphs, namely, sums of absolute coefficients of characteristic and distance polynomials.[18]

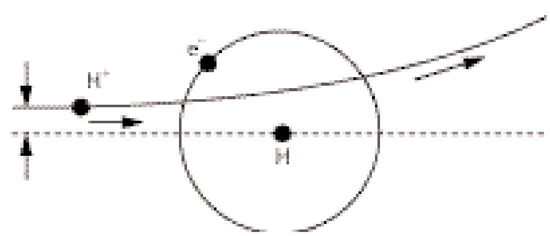
The transfer function (TF) designer module takes the HRG structures as its input and provides accentuated transfer function look-up tables for the in-house ray cast module for standard volume ray casting.



5 **Volume rendering of the metatorus with accentuated transfer functions. The top row (a-c) shows the volume rendering. In the bottom row (a-c), the bottom left represents the hue transfer function and the bottom right, the opacity transfer function. (a) With continuous hue and flat opacity transfer functions (reference); (b) with accentuated transfer functions according to Principle 1 ($\delta_h = 2/3\pi$, $\alpha = 0.02$, $\delta_o = 0.06$, $\omega_o = 0.04$); (c) with accentuated transfer functions according to Principle 2 ($\delta_h = 2/3\pi$, $\alpha = 0.02$, $\delta_o = 0.06$, $\omega_h = \omega_o = 0.03$).**



6 **AVS/Express module network for volume data mining.**



7 **The proton-hydrogen atom collision process.**

Application: Proton-hydrogen collision

To illustrate the feasibility of our methodology, we explored a large-scale 4D data set consisting of data for 61^3 volumes over 10^4 time steps for simulated intermediate-energy collisions of a proton and a hydrogen atom.[9] The simulation deals with a fundamental ion-atom collision problem and is very important in that the problem has a wide spectrum of applications such as nuclear fusion, material sciences, and radiology. We wanted to investigate how the positive charge of an incident proton affects the behavior of an electron around the target hydrogen atom (Figure 7). To this end, we obtained a comprehensible illustration of the collision by visualizing the 3D distorted electron density distribution.

The stationary electron density distribution around a hydrogen atom constitutes a completely layered structure of spherical isosurfaces. Therefore, without producing volume-rendered animation of the entire time sequence, we can easily identify the approximate timing of a collision by sampling snapshot volumes and searching the simplest structure among the constructed HRGs.

Figure 8 depicts three representative snapshots before, around, and after the collision: HRGs, isodensity surfaces extracted with a common target value throughout the entire time interval, and volume rendering instantaneously accentuated according to Principle 2. The three types of volume viewings allow us to understand more clearly the inner structures of the distorted electron density distribution.

Concluding remarks

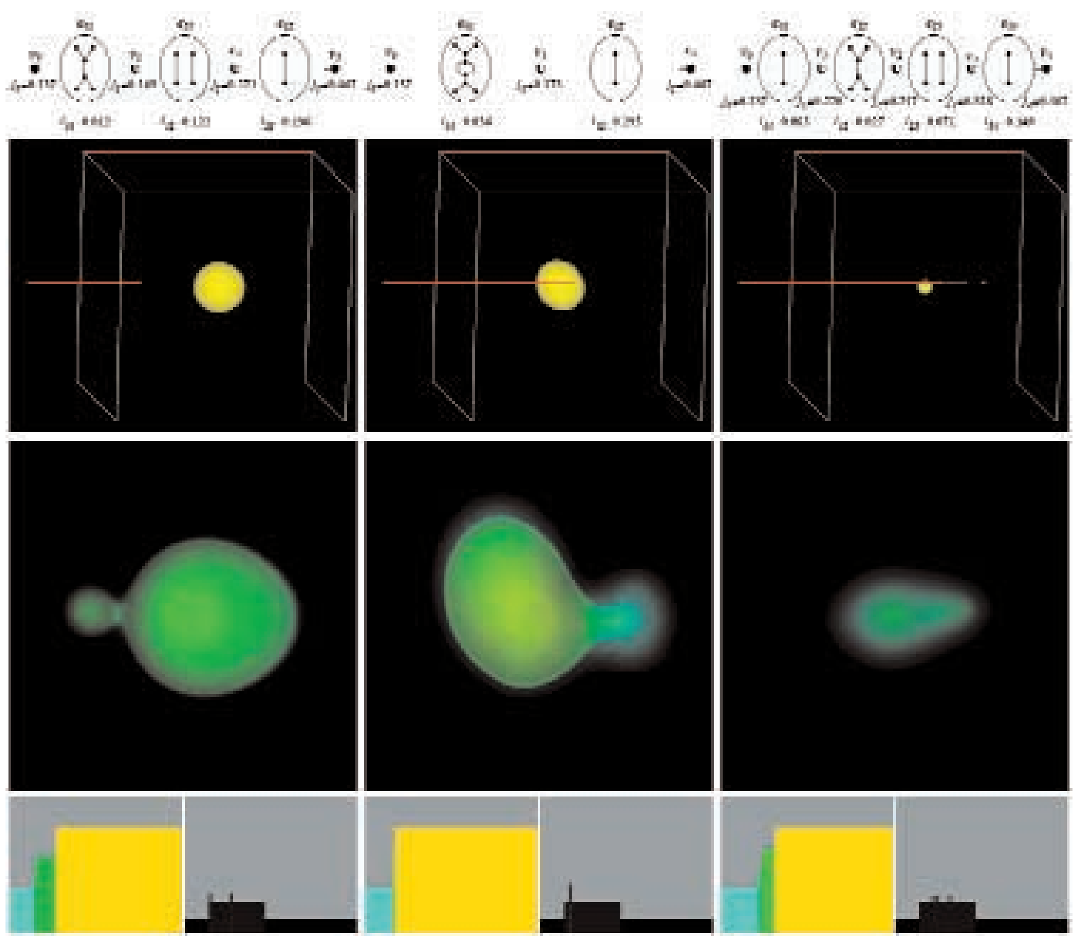
Combining sophisticated indirect/direct volume visualization with HRGs provides users with effective visual cues for discovering knowledge about the inner structure and complex behavior of volumetric objects. Our present methodology is quite general, we suggest potential benefits to recombining it with other approaches to volume rendering, such as splatting and cell projection. Moreover, the methodology is independent of the mesh type of an input volume data set, as far as a sequence of isosurfaces can be extracted from the data set.

On the other hand, the current implementation of the HRG constructor is very sensitive to the change in topology on volume boundaries, thus narrowing the analyzable field interval. An algorithm to extract a substantial subgraph from a given Reeb graph needs to be developed. In addition, it's crucial to choose an optimal direction for defining height-field functions to adjust the viewing direction. Omnidirectional (spherical) mapping looks like a promising solution to this problem, at the sacrifice of its high temporal complexity.

We view prioritizing CFVs to control the HRGs' level of detail as the key to improving the present methodology to a true volume data mining tool for various disciplines, including the medical and environmental sciences. Candidates for promising strategies include perturbation analysis of the direction for the height function axis and relaxation of topological equivalence of Reeb graphs. ■

Acknowledgements

We have benefited from continuous discussions with Xiaoyang Mao, Yoshihisa Shinagawa, Haruo Hosoya, Koji Koyamada, Sam Uselton, Karen Vierow, Yumi Yamaguchi, and research staff from Monolith, Inc. This work has been partially supported by the Ministry of Educa-



8 **Three types of volume data mining snapshots of a proton-hydrogen atom collision: HRG (first row); isodensity surface with the trajectory of the marching proton (second row); accentuated volume rendering (third row); and definitions of transfer functions (fourth row) with hue transfer on the left and opacity transfer on the right. (a) Before collision, (b) around collision, and (c) after collision.**

tion, Science, Sports, and Culture, the Japanese government under Grant-in-Aid for Scientific Research (C) 11680349.

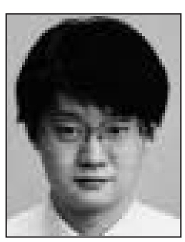
References

1. N. Ramakrishnan and A.Y. Grama, "Data Mining: From Serendipity to Science," *Computer*, Vol. 32, No. 8, Aug. 1999, pp. 34-37.
2. Y. Shinagawa and T.L. Kunii, "Constructing a Reeb Graph Automatically from Cross Sections," *IEEE Computer Graphics and Applications*, Vol. 11, No. 6, Nov./Dec. 1991, pp. 44-51.
3. S. Takahashi et al., "Algorithms for Extracting Correct Critical Points and Constructing Topological Graphs from Discrete Geographical Elevation Data," *Computer Graphics Forum*, Vol. 14, No. 3, Aug./Sep. 1995, pp. C181-C192.
4. I. Fujishiro, T. Azuma, and Y. Takeshima, "Automating Transfer Function Design for Comprehensible Volume Rendering Based on 3D Field Topology Analysis," *Proc. Visualization 99,* ACM Press, New York, 1999, pp. 467-470, 563.
5. Y. Shinagawa, T.L. Kunii, and Y.L. Kergosien, "Surface Coding Based on Morse Theory," *IEEE Computer Graphics and Applications*, Vol. 11, No. 5, Sep./Oct. 1991, pp. 66-78.
6. I.S. Kweon and T. Kanade, "Extracting Topographic Terrain Features from Elevation Maps," *CVGIP: Image Understanding*, Vol. 59, No. 2, Mar. 1994, pp. 171-182.
7. S. Takahashi, Y. Shinagawa, and T.L. Kunii, "A Feature-Based Approach for Smooth Surfaces," *Proc. ACM 4th Symp. on Solid Modeling and Applications*, ACM Press, New York, May 1997, pp. 97-110.
8. C.L. Bajaj, V. Pascucci, and D.R. Schikore, "The Contour Spectrum," *Proc. Visualization 97*, ACM Press, New York, 1997, pp. 167-173, 539.
9. I. Fujishiro et al., "Volumetric Data Exploration Using Interval Volume," *IEEE Trans. on Visualization and Computer Graphics*, Vol. 2, No. 2, Jun. 1996, pp. 144-155.
10. T. He et al., "Generation of Transfer Functions with Stochastic Search Techniques," *Proc. Visualization 96*, ACM Press, New York, 1996, pp. 227-234, 489.
11. J. Marks et al., "Design Galleries: A General Approach to Setting Parameters for Computer Graphics and Animation," *Computer Graphics Proc.*, Ann. Conf. Series, ACM Siggraph, ACM Press, New York, 1997, pp. 389-400.
12. S. Castro et al., "Transfer Function Specification for the Visualization of Medical Data," Tech. Report, TR-186-2-98-12, Vienna University of Technology, Mar. 1998, http://www.cg.tuwien.ac.at/research/TR/98/TR-186-2-98-12Abstract.html.
13. G. Kindlmann and J.W. Durkin, "Semi-Automatic Generation of Transfer Functions for Direct Volume Rendering," *Proc. Symp. on Volume Visualization*, ACM Siggraph, ACM Press, New York, 1998, pp. 79-86, 170.
14. S. Fang, T. Biddlecome, and M. Tuceryan, "Image-Based Transfer Function Design for Data Exploration," *Proc. Visualization 98*, ACM Press, New York, 1998, pp. 319-326, 546.
15. H.D. Lord, "Improving the Application Visualization Development Process with Modular Visualization Environments," *ACM Computer Graphics*, Vol. 29, No. 2, May 1995, pp. 10-12.
16. G. Taubin, "A Signal Processing Approach to Fair Surface Design," *Computer Graphics Proc.*, Ann. Conf. Series, ACM Siggraph, ACM Press, New York, 1995, pp.351-358.
17. J.L. Pfaltz, "Surface Networks," *Geographical Analysis*, Vol. 8, 1976, pp. 77-93.
18. L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, San Diego, Calif., 1976.

Issei Fujishiro *is a professor of information sciences at Ochanomizu University and an invited researcher at the Research Organization for Information Science and Technology (RIST). His research interests include volume visualization and graphics, information visualization design, and immersive virtual environments. He received his ME in information sciences and electronics from the University of Tsukuba, Japan, in 1985 and his DSc in information sciences from the University of Tokyo in 1988. He is serving on the editorial boards for* IEEE Transactions on Visualization and Computer Graphics *and* Transactions of the Information Processing Society of Japan *(IPSJ).*

Taeko Azuma *is currently working for the Consulting Solution Business Unit of Hewlett-Packard, Japan. She received her BS and MS in information sciences from Ochanomizu University in 1998 and 2000, respectively. Her research interests include volume modeling and comprehensible rendering.*

Yuriko Takeshima *is a research associate in the Graduate School of Humanities and Sciences, Ochanomizu University, and an invited researcher at RIST. She received her BS and MS in information sciences and her PhD in the Doctoral Research Course in Human Culture from Ochanomizu University in 1994, 1996, and 1999, respectively. Her research interests include indirect volume visualization and multivariate data visualization. She is a member of the IEEE Computer Society, ACM, ACM Siggraph, and IPSJ.*



Shigeo Takahashi *is an associate professor of the Computer Center at Gunma University. He received his BS, MS, and PhD degrees in information science from the University of Tokyo in 1992, 1994, and 1997, respectively. His research interests include computer graphics and geometric modeling. He is a member of the IEEE Computer Society, ACM, IPSJ, and the Institute of Electronics, Information, and Communications Engineers.*

Contact Fujishiro at the Dept. of Information Sciences, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan, e-mail fuji@is.ocha.ac.jp.