Verity Visualization: Visual Mappings

Craig M. Wittenbrink
Alex T. Pang
Suresh Lodha*

UCSC-CRL-95-48
October 11, 1995

Baskin Center for
Computer Engineering & Information Sciences
University of California, Santa Cruz
Santa Cruz, CA 95064 USA

ABSTRACT

Visualized data often has dubious origins. One way to define data lineage is by describing the uncertainty. In addition, different forms of uncertainty and errors are also introduced as the data is derived, transformed, interpolated, and finally rendered. In the absence of integrated presentation of data and its associated uncertainty, the analysis of the visualization is incomplete at best and often leads to inaccurate or incorrect conclusions. This paper presents several techniques of presenting data together with uncertainty. The idea behind these techniques can be applied to both spatial (e.g. surface) and temporal (i.e. animation) domains. We describe these techniques of representing the truths about the data as verity visualization. The same techniques can also be used to make the users aware of the data quality or to emphasize and draw their attention to the uncertainty.

Keywords: data quality, uncertainty glyphs, fat surfaces, perturbations, oscillations.

*Partially supported by a grant from NSF IRI-9423881 and ONR grant N00014-92-J-1807

1 Introduction

With few exceptions, most of the visualization work done to date have ignored or isolated the presentation of uncertainty from the data. Part of the reason for this practice is the inherent difficulty in defining, characterizing, and controlling the introduction of uncertainty in the visualization pipeline (figure 2.1). Another difficulty is the absence of methods that integrate the presentation of data together with uncertainty. Finally, there is also a need for a framework to evaluate the effectiveness of these verity visualization methods. This paper focuses on the problem of visually mapping data and uncertainty together into a holistic view.

From one perspective, one might consider adding uncertainty parameters as additional dimensions or fields to visualize using existing surface, volume, flow, and multi-dimensional visualization methods. In fact, we do start with existing methods. However, even with the simple task of designing glyphs or icons that incorporate uncertainty information [1, 2, 3], the process is sometimes counter-intuitive. For example, while a glyph may appear appropriate by itself, the user's perception of the glyph may be different when a group of them is presented in various scales and locations. Thus, while some of the methods we have examined are not necessarily new, they must be able to render and convey the data in complete accordance with the facts. We call this *verity visualization* since the word verity (according to Webster) suggests the quality or state of being true or real. While this has been recognized and is often stated as a worthy goal in scientific visualization (e.g. in the IEEE Visualization discussion on *How to Lie with Visualization* and the NCGIA initiative on *Visualization of Spatial Data Quality* [4]), it has rarely been pursued or realized. This paper presents some methods that represent significant steps toward achieving this goal.

2 Uncertainty

2.1 What is Uncertainty?

We define uncertainty as statistical variation or spread, error, and minimum-maximum ranges. NIST has written a standards report on uncertainty, which includes operator error [5], but for the discussion in this paper we consider three types of uncertainty: *statistical* – either given by the estimated mean and standard deviation, which can be used to calculate a confidence interval, or an actual distribution of the data; *error* – a difference, or an absolute valued error among estimates of the data, or between a known correct datum and an estimate; and *range* – an interval in which the data must exist, but which cannot be quantified into either the statistical or error definitions. Note that the term data quality has an inverse relationship with data uncertainty [6] and hence can also take advantage of the techniques presented in this paper.

2.2 Sources of Uncertainty

In order to understand what is overlooked in visualization, we quickly review the sources of uncertainty, errors, and ranges within data. Figure 2.1 illustrates the three major blocks in a visualization pipeline leading to the analysis of the visualization output. It is clear that different forms of uncertainty are introduced into the pipeline as data are acquired, transformed, and visualized. Starting with the data acquisition stage, one will note that nearly all data sets, whether from instrument measurements, numerical models, or data entry have a statistical variation [7]. With instruments, there is an experimental variability whether the measurements are taken by a machine or by a scientist. The more times the measurement is taken, the more confident the measurement. But there will be a statistical variation in these measurements. The same is true for data from numerical models and human observations or inputs. In numerical modeling, the model and its parameters have been decided by a domain specialist, and is inherently a simplification (e.g. linearization of a nonlinear system) of the system being modeled. In addition to model simplification and sensitivity of these models to input parameters, numerical calculations performed on these models also introduce errors due to the integration algorithms and the limited precision of the computing machinery. Likewise, there is variability in human observations both in terms of difference in perception among individuals and also to slight differences when asked to perform a task repeatedly.

2.3 Uncertainty in Visualization

As can be seen in figure 2.1, derived uncertainty is introduced in the transformation or sec-

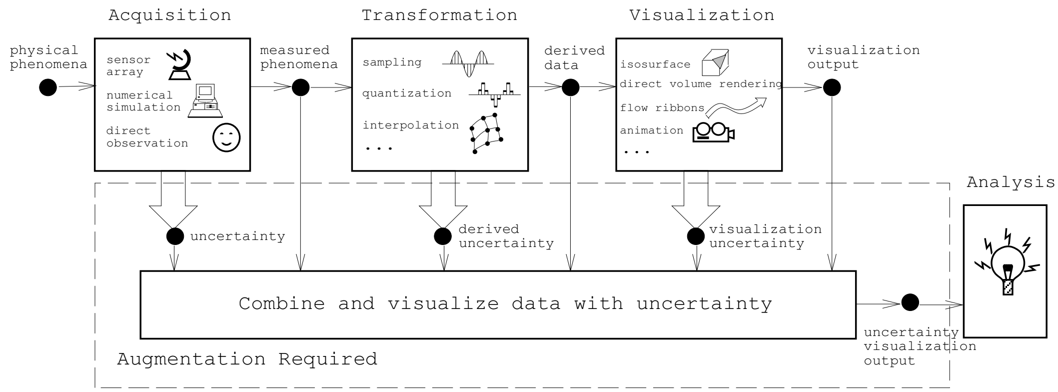

Figure 2.1: This visualization pipeline shows measurement uncertainty, derived uncertainty, and visualization uncertainty.

ond stage of the visualization pipeline. What is more interesting and perhaps not self evident is that uncertainty is also introduced in the visualization stage itself. Within the area rendering with radiosity, there has been some recent work in controlling the errors introduced in the rendering process [8, 9, 10]. As these researchers also pointed out, the rendering process introduces uncertainty arising from the data collection process, algorithmic errors, and computational accuracy and precision.

Aside from radiosity, other rendering and visualization methods also suffer from unintentional and perhaps unavoidable errors introduced during the visualization process. For example, while the holes arising from ambiguities in the marching cubes algorithm [11] have been fixed [12], the iso-surfaces are obtained using interpolation and may not reconstruct the original surface. The same is true for flow visualization methods, where implementors are faced with decisions on which integration algorithm to use. Surface modeling and animation are not immune. In surface interpolation a variety of tradeoffs exist in performance and results, and there is no ideal surface in many cases because of the many free parameters available [13]. In many cases the data that are to be interpolated have numerous errors, and may even lack topology information [14]. In animation, the process of creating the key frames is error prone. The in-betweening to fill in frames between the key frames is analogous to surface interpolation, and though no method is correct, there are many methods available, and all of them will result in slight variations.

3 Existing Methods of Visualizing Uncertainty

Many researchers are fully aware of the uncertainty in their data usually in the form of errors. These are usually displayed using some straightforward method such as side by side comparison or differencing. For example, [9] used line plots to render uncertainty, [8] used difference images, and [10] used norms for the entire image. In surface interpolation, pseudo-coloring of the surface curvature or other properties of the surface is used [15].

In geographic and information systems, researchers are aware of the statistical variation, and have been more creative. However, they use essentially multivalued visualization methods, and simply add uncertainty as another parameter into the picture. For example, GIS researchers used the color of the areas on a map to represent the uncertainty of the data at that point on the map. They have assumed that the variety of techniques available is fixed, and wish to simply use available solutions from the visualization community.

New techniques are being developed for higher order data such as tensors, for new hardware features such as texture mapping for flow visualization, and for adding more and more variables into existing methods such as streamlines which result in stream balls [16]. Some newer approaches include animation for the display of uncertainty in fuzzily classified regions [17]. With few exceptions, most of the existing methods for visualizing uncertainty rely on the *overloading* approach where uncertainty parameters are treated as additional data fields to be mapped to visual cues. This has the disadvantage of contention between data and uncertainty information for the visual cues. The approach that we are advocating is called the *verity visualization* approach where new and/or modification of existing techniques are used to integrate the display of both data and uncertainty in the same picture without using overloading. We believe that these techniques will help the scientists, graphics users, and lay people doing visualization. One example of our work is the development of a new type of vector glyph which shows statistical variation, error, or range in both the magnitude and bearing [3]. Another one uses iterated function systems to indicate the level of uncertainty in surface interpolation [18].

We have done a classification of uncertainty visualization techniques, and concluded that only the scalar low density plot has been adequately explored, where the uncertainty may be shown with economy using Tukey's box plots [19], Tufte's quartile plots [20] and/or Cleveland's framed rectangles [21]. What we demonstrate in the following section are new methods for displaying higher dimensional uncertainty (e.g. a vector of uncertainty parameters) in surfaces and in animation applications.

4 Visual Mappings of Data with Uncertainty

We present four different methods, representative of the verity visualization approach, which presents data and uncertainty in an integrated fashion. These methods are: uncertainty glyphs, fat surfaces, surface perturbations, and oscillations. Although these methods imply the existence of some surface, we will show that they can also be applied to the visualization of uncertainty in animation algorithms.

Uncertainty glyphs: Glyphs or icons are graphics objects that encode information through their shape, color, size, and other attributes. Uncertainty glyphs are probes which can be placed in a graphic to indicate the confidence interval, error, or range. Examples of uncertainty glyphs for vector fields were presented in [3] and included both the uncertainty in direction and magnitude of the vector. The challenging aspects of uncertainty glyph design are in the design of their shapes, density and placement, and scaling.

Fat surfaces: These are surfaces or envelopes which show the range of possible values in the data. They are most appropriate for uncertainty represented by a range of min/max values.

Perturbations: The idea here is to represent uncertainty as randomized surface roughness. These perturbed surfaces give an indication of the location and degree of uncertainty in the data.

Oscillations: This is an alternative way of presenting min/max values. Instead of using fat surfaces, a single surface is made to oscillate between the range of possible values. An extra degree of freedom with oscillation is to map the duration of the surface at a particular position to the likelihood of the data value at that position.

We now illustrate how these methods can be used to visualizing uncertainty in surface modeling and animation.

4.1 Visualizing uncertainty in surface interpolations

As an illustration for the four methods of visualizing uncertainty in surface interpolations, consider the errors or differences between two interpolation methods: bilinear and multi-quadric. Figure 4.1 shows the bilinearly interpolated surface and figure 4.2 shows the surface obtained through multi-quadric. As mentioned, traditional approaches at visualizing the differences between the two interpolation methods include side by side comparison, difference images (figure 4.3), pseudo coloring the differences (figure 4.4), and transparency. On the other hand, using our proposed methods, they appear as figure 4.5 with line glyphs or figure 4.6 with ellipsoidal glyphs. The glyph shapes are simple in this case as the the uncertainty parameter is simply the magnitude of the difference. With fat surfaces, figure 4.7 indicates the distance between the two interpolated surfaces. The "fat" parameter has been scaled up to emphasize the difference between the two surfaces. To indicate the regions where the most variation occurs, we use surface perturbations as illustrated in figure 4.8. As described, oscillations can be used to indicate the location and magnitude of the differences between the two surfaces. (These can be seen in the accompanying video). The point to note with all of these figures is that the uncertainty information is combined with the rendering itself.

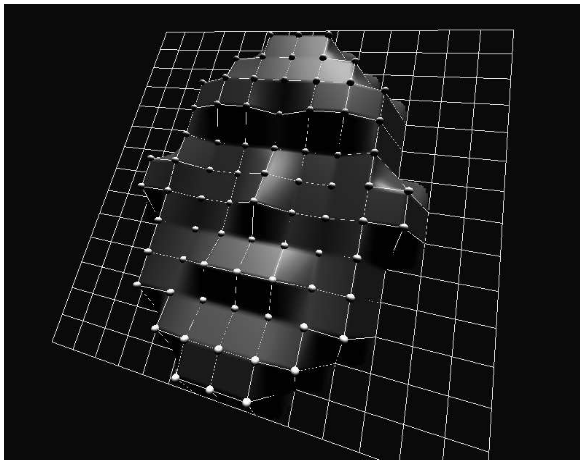

Figure 4.1: Bilinearly interpolated surface.

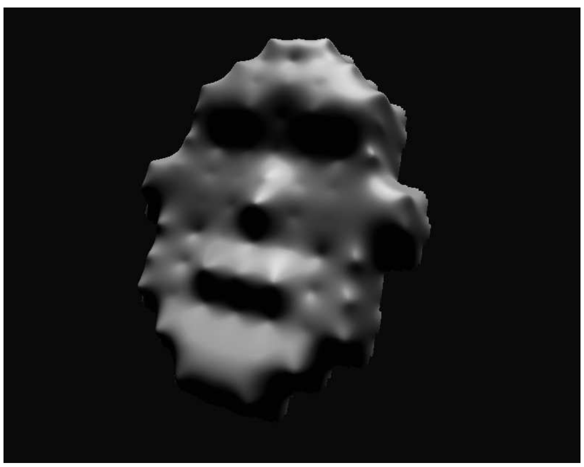

Figure 4.2: Surface obtained using multi-quadric interpolation.

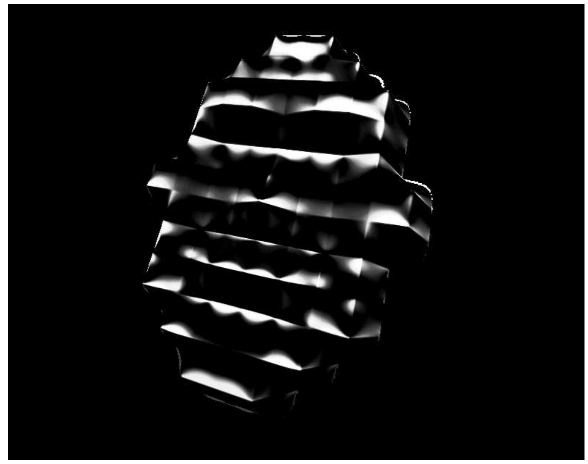

Figure 4.3: Difference image of figures 4.1 and 4.2.

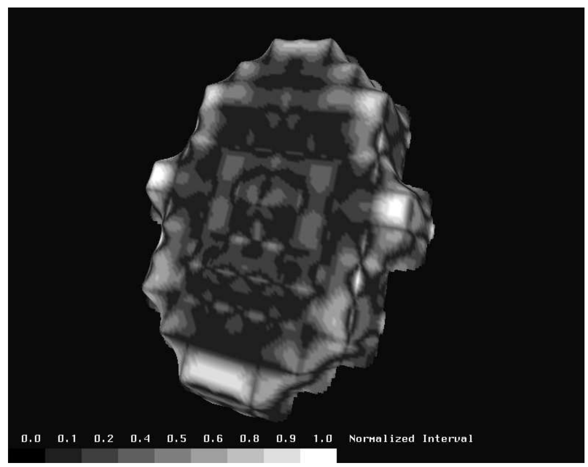

Figure 4.4: Pseudo colored difference on surface of image.

4.2 Visualizing uncertainty in animations

A popular method of animation is by specifying key-frames and generating the in-between frames using interpolation. This method is usually used in character animation and more recently in morphing. Depending on the interpolation method selected, the animation paths may vary slightly. In this section, we use a simple animation over a 2D M-shaped path to illustrate how the verity visualization methods can be used to highlight the differences between a linear and a cubic interpolation method for generating the in-between frames.

As with the surface interpolation example, the uncertainty parameter here is an error term between the position/path of the in-between frames. Figure 4.9 shows how simple line glyphs indicate the paths and positions of linear (red) interpolation versus cubic (green) interpolation. Alternatively, sphere based glyphs can also be used (figure 4.10). The equivalent of fat surfaces in animation is to simultaneously animate the balls (or actors) using both paths. Figure 4.11 is a snapshot showing how random path perturbation (constrained or controlled by the difference) can be used instead. Finally, we use motion blurring to indicate the variation between the two paths in figure 4.12.

5 Evaluation

So far, we have presented some verity visualization methods. We are in the process of adding other methods and extending the application base. More importantly, we are also in

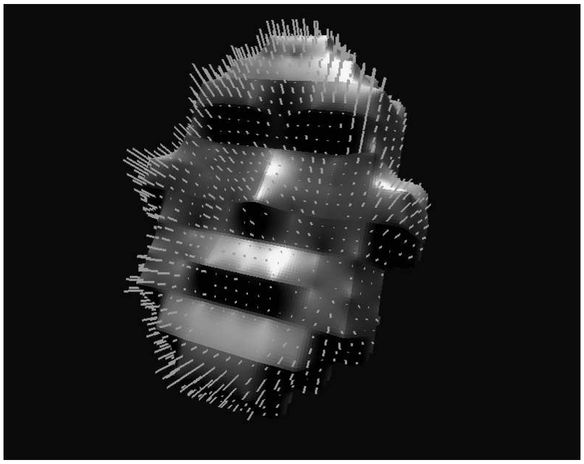

Figure 4.5: Bilinear interpolated surface with difference between bilinear and multi-quadric shown with line glyphs.

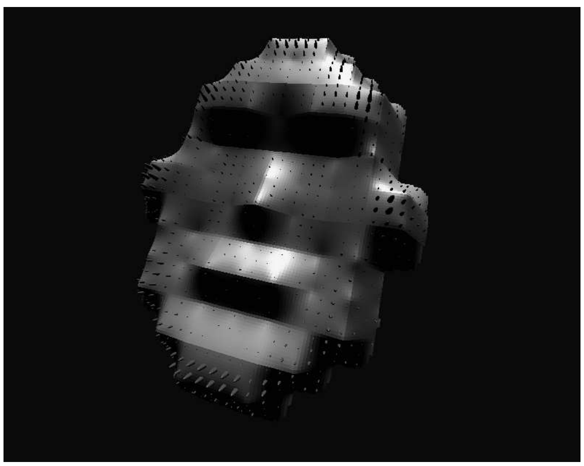

Figure 4.6: Bilinear interpolated surface with difference between bilinear and multi-quadric shown with ellipsoidal glyphs.

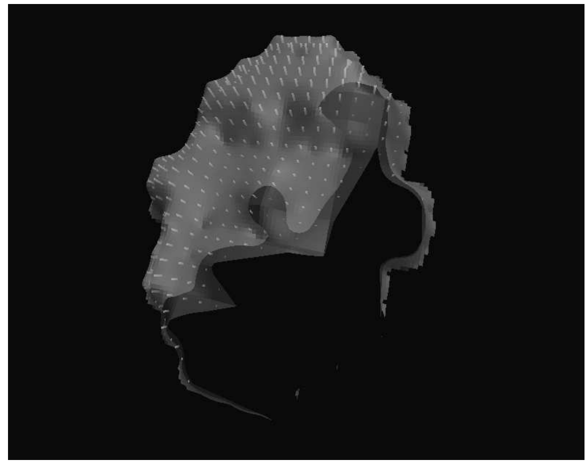

Figure 4.7: Fat surfaces from the bilinear and multi-quadric interpolations with line glyphs. The clipping plane Is used to show the exaggerated difference.

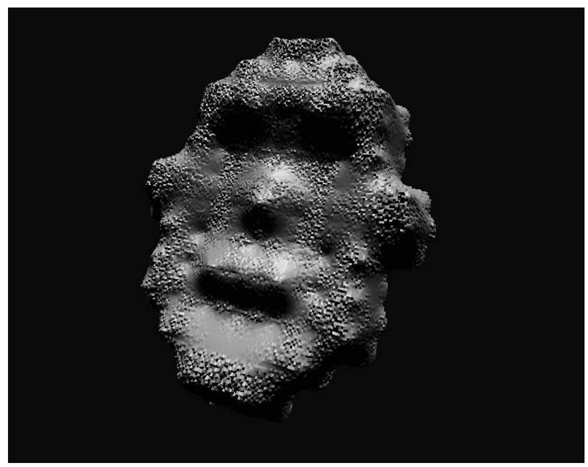

Figure 4.8: Random displacement mapping scaled to the difference to give an indication of areas with a greater deviation.

the process of evaluating the effectiveness of these new methods. Two approaches are being taken in this effort. One, the *quantitative* approach provides a domain independent measure. Examples include those suggested by Tufte [20, 2]: data-ink maximization, clutter and moire pattern minimization, and multi-functionality of graphic elements. Two, the *qualitative* approach provides a more subjective measure of the methods. The measure may vary among different application domains. In addition, it relates to perceptual issues, the ability of the user to correlate data and uncertainty from the presentation, training time to understand the new presentation, relative improvements over existing methods, and any changes in the conclusions drawn from the presentation, etc. The results from this study will be the subject of another paper.

6 Conclusions

In this paper, we presented some verity visualization methods (uncertainty glyphs fat surfaces, perturbations, and oscillations) and applied them to surface interpolation and animation applications. The resulting visualizations of data and

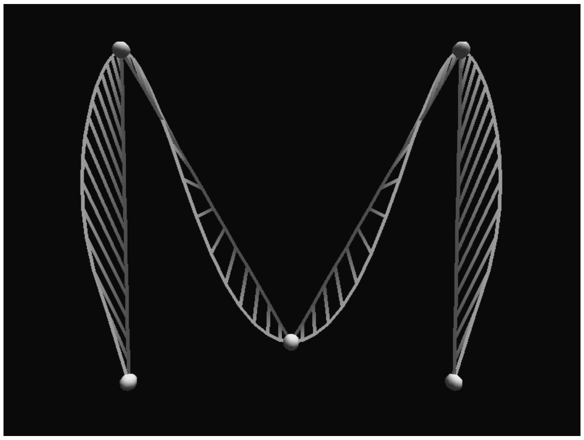

Figure 4.9: Animation with line glyphs and paths.

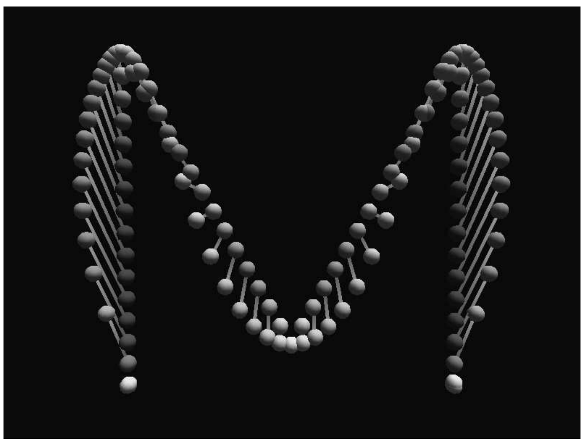

Figure 4.10: Animation with sphere glyphs and paths.

uncertainty are integrated and present an accurate depiction to the user. We believe that these methods will prove valuable to people who need to make informed decisions based on imperfect data.

7 Acknowledgements

We wish to thank Michael Clifton and Bob Sheehan for their inputs and programming efforts. We would also like to thank Jane Wilhelms and Allen Van Gelder for their assistance and feedback.

References

[1] John W. Tukey. *The Collected Works of John W. Tukey: Volume V: Graphics,: 1965-1985.* Wadsworth and Brooks, 1984.

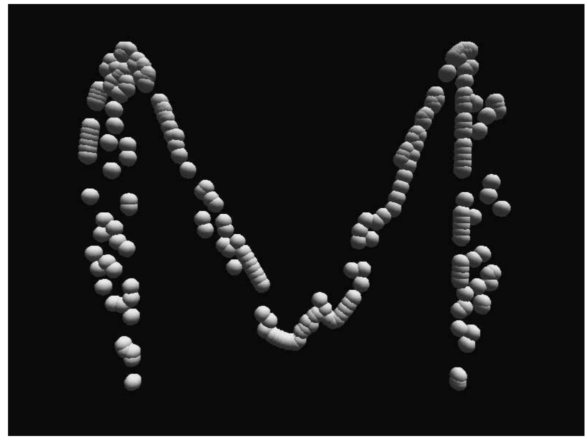

Figure 4.11: Animation with random perturbation to indicate uncertainty.

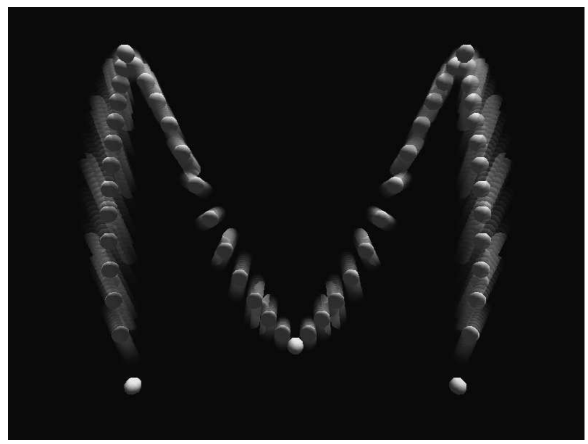

Figure 4.12: Animation with motion blurring to indicate uncertainty.

[2] Edward R. Tufte. *Envisioning Information.* Graphics Press, 1990.

[3] Craig M. Wittenbrink, Elijah Saxon, Jeff J. Furman, Alex T. Pang, and Suresh Lodha. Glyphs for visualizing uncertainty in environmental vector fields. In *SPIE & IS&T Conference Proceedings on Electronic Imaging: Visual Data Exploration and Analysis.* SPIE, Feb. 7–10 1995.

[4] M. Kate Beard, Barbara P. Buttenfield, and Sarah B. Clapham. NCGIA research initiative 7: Visualization of spatial data quality. Technical Paper 91-26, NCGIA, October 1991.

[5] Barry N. Taylor and Chris E. Kuyatt. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. Techni-

cal report, NIST Technical Note 1297, January 1993.

[6] Alex Pang, Jeff Furman, and Wendell Nuss. Data quality issues in visualization. In *SPIE Vol. 2178 Visual Data Exploration and Analysis*, pages 12–23. SPIE, February 1994.

[7] Christopher Chatfield. *Statistics for Technology, A Course In Applied Statistics*. Chapman and Hall, third edition, 1983.

[8] Ned Greene and Michael Kass. Error-bounded antialiased rendering of complex environments. In *Proceedings of SIGGRAPH 94*, pages 59–66, Orlando, FL, Jul. 24-29 1994. ACM SIGGRAPH.

[9] Dani Lischinski, Brian Smits, and Donald P. Greenberg. Bounds and error estimates for radiosity. In *Proceedings of SIGGRAPH 94*, pages 67–74, Orlando, FL, Jul. 24-29 1994. ACM SIGGRAPH.

[10] James Arvo, Kenneth Torrance, and Brian Smits. A framework for the analysis of error in global illumination algorithms. In *Proceedings of SIGGRAPH 94*, pages 75–84, Orlando, FL, Jul. 24-29 1994. ACM SIGGRAPH.

[11] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21(4):163 – 169, 1987.

[12] Allen Van Gelder and Jane Wilhelms. Topological considerations in isosurface generation. *ACM Transactions on Graphics*, 13(4), October 1994. Also UCSC technical report UCSC-CRL-90-14.

[13] Gerald Farin. *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide*. Academic Press, 88.

[14] Hugues Hoppe, Tony DeRose, Tom Duchamp, Mark Halstead, Hugert Jin, John McDonald, Jean Schweitzer, and Werner Stuetzle. Piecewise smooth surface reconstruction. In *Proceedings of SIGGRAPH 94*, pages 295–303, Orlando, FL, Jul. 24-29 1994. ACM SIGGRAPH.

[15] Hans Hagen, Stefanie Hahmann, Thomas Schreiber, Yasuo Nakajima, Burkard Wordenweber, and Petra Hollemann-Grundstedt. Surface interrogation algorithms. *IEEE Computer Graphics and Applications*, 12(5):53–60, September 1992.

[16] Manfred Brill, Hans Hagen, Hans-Christian Rodrian, Wladimir Djatschin, and Stanislav V. Klimenko. Streamball techniques for flow visualization. In *Proceedings: Visualization '94*, pages 225–231. IEEE Computer Society, 1994.

[17] Nahum D. Gershon. Visualization of fuzzy data using generalized animation. In *Proceedings of Visualization 92*, pages 268–273. IEEE Computer Society Press, October 1992.

[18] Craig M. Wittenbrink. Ifs fractal interpolation for 2d and 3d visualization. Technical report, University of California, Santa Cruz, Santa Cruz, CA, 1995. submitted to Visualization'95.

[19] John W. Tukey. *Exploratory Data Analysis*. Addison Wesley, 1977.

[20] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 1983.

[21] William S. Cleveland. *The Elements of Graphing Data*. Wadsworth, 1985.