Data Visualization Using Automatic, Perceptually-Motivated Shapes

Christopher D. Shaw
University of Regina *

David S. Ebert, James M. Kukla, Amen Zwa, Ian Soboroff
U. of Maryland Baltimore County †

D. Aaron Roberts
NASA Goddard Space Flight Center ‡

ABSTRACT

This paper describes a new technique for the multi-dimensional visualization of data through automatic procedural generation of glyph shapes based on mathematical functions. Our glyph-based Stereoscopic Field Analyzer (SFA) system allows the visualization of both regular and irregular grids of volumetric data. SFA uses a glyph's location, 3D size, color and opacity to encode up to 8 attributes of scalar data per glyph. We have extended SFA's capabilities to explore shape variation as a visualization attribute. We opted for a procedural approach, which allows flexibility, data abstraction, and freedom from specification of detailed shapes. Superquadrics are a natural choice to satisfy our goal of automatic and comprehensible mapping of data to shape. For our initial implementation we have chosen superellipses. We parameterize superquadrics to allow continuous control over the "roundness" or "pointiness" of the shape in the two major planes which intersect to form the shape, allowing a very simple, intuitive, abstract schema of shape specification.

Keywords: Glyphs, Volume Visualization, Information Visualization, Superquadrics, Two-Handed Input

1. INTRODUCTION

The simultaneous visualization of multi-dimensional data is a difficult task. The goal is not only the display of multi-dimensional data, but the *comprehensible display* of multi-dimensional data. Glyph, or iconic, visualization is an attempt to encode more information in a comprehensible format, allowing multiple values to be encoded in the parameters of the glyphs.[1] The shape, color, transparency, orientation, etc., of the glyph can be used to visualize data values. Glyph rendering[1,2] is an extension to the use of glyphs and icons in numerous fields, including cartography, logic, and pictorial information systems.

In previous work, we explored the usefulness of stereo-viewing and two-handed interaction to increase the perceptual cues in glyph-based visualization. The Stereoscopic Field Analyzer (SFA)[3] allows the visualization of both regular and irregular grids of volumetric data. SFA combines glyph-based volume rendering with a two-handed minimally-immersive interaction metaphor to provide interactive visualization, manipulation, and exploration of multivariate, volumetric data. SFA uses a glyph's location, 3D size, color and opacity to encode up to 8 attributes of scalar data per glyph. These attributes are used when a vector visualization is not appropriate, such as when displaying temperature and pressure at each glyph. We are extending this work to combine glyph rendering with other visually salient features to increase the number of data dimensions simultaneously viewable.

2. BACKGROUND

We chose to explore shape variation based on its priority in human perception. Cleveland[4] cites experimental evidence that shows the most accurate method to visually decode a quantitative variable in 2D is to display position along a scale. This is followed in decreasing order of accuracy by interval length, slope angle, area, volume, and color. Bertin offers a similar hierarchy in his treatise on thematic cartography.[5]

SFA employs glyph position in 3D, 3D scale (corresponding to Cleveland's length, area and volume) and color. Except in the vector case, SFA cannot use slope angle to encode information in a glyph because the 3D volume containing the glyphs can be turned to an arbitrary orientation. The use of slope angle to encode a scalar value depends on a consistent frame of reference,

*Dept of Computer Science, U. of Regina, Regina, Saskatchewan, Canada S4S 0A2, Ph: 306-585-4071, email: cdshaw@cs.uregina.ca
†CSEE Dept. UMBC, 1000 Hilltop Circle , Baltimore, MD 21250, Phone: (410)455-3541, email: ebert@cs.umbc.edu
‡NASA Goddard Space Flight Center Mailstop 692.0, Greenbelt, MD 20771, roberts@vayu.gsfc.nasa.gov

but SFA allows the frame of reference to be changed at any time. For example, looking at a tilted glyph from the front and from the back will visually offer opposite readings because the glyph is visually reversed.

Therefore, the next opportunity for encoding a scalar value is shape. One of the most difficult problems in glyph visualization is the design of meaningful glyphs. Glyph shape variation must be able to convey changes in associated data values in a comprehensible manner.[1] This difficulty is sometimes avoided by adopting a single base shape and scaling it non-uniformly in 3 dimensions. However, the lack of a more general shape interpolation method has precluded the use of shape beyond the signification of categorical values.[5] This paper describes a new system for the procedural generation of glyph shapes for glyph-based volumetric visualization[6] using superquadrics.[7]

2.1. Procedural shape visualization using superquadrics

Because of the need for meaningful glyph design and the complexity of the problem, we opted for a procedural approach, which allows flexibility, data abstraction, and freedom from specification of detailed shapes.[8] Procedural techniques provide enormous flexibility to the user by allowing the shape to be controlled by high-level control parameters. These allow the user to change the glyph shape from a more directorial, indirect aspect, where he or she is unburdened from the full explicit specification of detailed shapes. Our goal for glyph design was to allow the automatic mapping of data to shape in a comprehensible, easily controllable manner. Superquadrics are a natural choice to satisfy this goal. Superquadrics, first introduced to computer graphics by Barr,[7] are extensions of quadric surfaces where the trigonometric terms are each raised to exponents. Superquadrics come in four main families: hyperboloid of one sheet, hyperboloid of two sheets, ellipsoid, and toroid. For our initial implementation we have chosen superellipses due to their familiarity, but the system can be easily extended to use other types of superquadrics as well as combinations of types. For example, supertoroids could be used for negative values and superellipsoids for positive values.

In the case the of superellipsoids, the trigonometric terms are assigned exponents as follows:

$$\underline{x}(\eta,\omega) = \begin{bmatrix} a_1 \quad \cos^{\epsilon 1}\eta \quad \cos^{\epsilon 2}\omega \\ a_2 \quad \cos^{\epsilon 1}\eta \quad \cos^{\epsilon 2}\omega \\ a_3 \quad \sin^{\epsilon 1}\eta \end{bmatrix}, \quad \begin{matrix} -\pi/2 \leq \eta \leq \pi/2 \\ -\pi \leq \omega < pi \end{matrix}$$

These exponents allow continuous control over the characteristics (in some sense the "roundness" or "pointiness") of the shape in the two major planes which intersect to form the shape, allowing a very simple, intuitive, abstract schema of shape specification. For example, as Barr states, $\epsilon 1 < 1$ and $\epsilon 2 < 1$ produces cuboid shapes, $\epsilon 1 < 1$ and $\epsilon 2 \sim 1$ produces cylindroid shapes, $\epsilon 1 > 2$ or $\epsilon 2 > 2$ produces pinched shapes while $\epsilon 1 = 2$ or $\epsilon 2 = 2$ produces faceted shapes. As can be seen in Figure 1, varying the exponents achieves smooth, understandable transitions in shape. Therefore, mapping data values to the exponents provides not only a continuous, automatic control over the shape's overall flavor, but a comprehensible shape mapping as well.

3. PERCEPTUALLY-BASED MAPPING OF SHAPE ATTRIBUTES

By using superquadrics, we can provide the appropriate shape visual cues for discerning data dimensions mapped to glyph shape while not distracting from the cognition of global data patterns.

Glyph shape is a valuable visualization component because of the human visual system's pre-attentive ability to discern shape. Shapes can be distinguished at the pre-attentive stage[9] using curvature information of the 2D silhouette contour and, for 3D objects, curvature information from surface shading. Unlike an arbitrary collection of icons, curvature has a visual order, since a surface of higher curvature looks more jagged than a surface of low curvature. Therefore, generating glyph shapes by maintaining control of their curvature will maintain a visual order. This allows us to generate a range of glyphs which interpolate between extremes of curvature, thereby allowing the user to read scalar values from the glyph's shape. Pre-attentive shape recognition allows quick analysis of shapes and provides useful dimensions for comprehensible visualization.

Our use of glyphs is related to the idea of marks as the most primitive component that can encode useful information.[5] Senay points out that shape, size, texture, orientation, transparency, hue, saturation, brightness, and transparency are retinal properties of marks that can encode information.[10,11] Bertin has studied the use of marks for two-dimensional thematic maps and gives examples of how shape can be misused in the rendering of these maps.[5] In his examples, shapes are used to represent purely categorical data and, for this reason, he uses a small collection of distinct icons such as star, cross, square, circle, triangle, and so on. Because each individual shape does not have any inherent meaning, the reader is forced to continually look up the shape's meaning in the map legend. The main difficulty is that a collection of arbitrary icons does not have any necessary visual order, and so any assignment of shape to meaning is equivalent.

To produce understandable, intuitive shapes, we are relying on the ability of superquadrics to create graphically distinct,[10,11] yet related shapes. We are encoding two data dimensions to glyph shape in a manner that allows the easy separation of the shape characteristics.

Since size and spatial location are more significant cues than shape, the importance mapping of data values should be done in a corresponding order. In decreasing order of data importance, data values were mapped to location, size, color, and shape. In our experience, shape is very useful for local area comparisons among glyphs: seeing local patterns, rates of change, outliers, anomalies.

We chose to map either one independent variable to both glyph exponents or two related variables to each glyph exponent to ensure the understandability of the shapes.

4. IMPLEMENTATION AND RESULTS

To exemplify the use of superquadric surfaces as a means of parametric glyph design, we implemented a system for visualizing large data sets as a shape previewer for SFA. The 2D interface to the system is written in C++ and Motif, while the rendering code is written in C and OpenGL.

4.1. Interface Design

SFA allows the user to specify a number of control parameters through a glyph mapping dialog. This gives the user the ability to select any of the input (data) parameters and map it to any of the output (display) parameters (glyph attributes). These mappings are one-to-many to eliminate data collisions while determining the character of an output parameter. The data parameters may be linearly mapped, inversely mapped, or procedurally mapped to shape attributes, providing the flexibility to choose alternate relations of the same parameters. In addition to location, seven current glyph attributes are provided: the two exponents of the superquadric shape, size in the X, Y, and Z directions, color and opacity. There is a special output parameter "- None -" designed to allow the user to disable the display of one or more attributes and simplify the display.

4.2. Glyph Generation

During glyph generation the value of each of the seven parameters is chosen. Each data variable is normalized to the range [0,1] and mapped into the domain of the corresponding glyph attribute. The choices for color and opacity are mapped onto lookup tables, and the size scale values are mapped between a pair of minimum and maximum size scale values.

The scalar value assigned to a superquadric exponent is similarly mapped using a table lookup procedure. This map specifies discrete data ranges (buckets) of the scalar value and the corresponding exponent value for that data range. The data range for each bucket can be generated based on a linear, inverse linear, or procedurally generated function. The lookup table also allows for on-demand creation of shapes, which can significantly decrease memory usage at run-time at only a slight cost during the shape generation phase. Shape generation need only occur upon changes in the mapping function between scalar data value and superquadric exponent. Once a satisfactory set of mappings is achieved, the user may explore the data set without the added cost of regenerating each glyph in the scene. The minimum and maximum ranges for each parameter are user dependent; we have chosen [0,2] for scaling, [0,255] for color and opacity maps, and [0,5] for the exponent parameters.

In simultaneously visualizing thousands of shapes, it becomes apparent that not all regions of the superquadric exponent domain map to equally rich groups of shapes. For example, the difference in shape between exponent values 3 and 4 is far less than that between 0 and 1. Therefore there are some particularly rich regions of shape that should be taken into consideration and exploited by the mapping routine. In the case that the exponent map is hand-made, the creator of the map should take this into consideration. We are experimenting with coding methods that automatically take this perceptual variability into account.

4.3. Information Visualization Results

We have applied procedurally-generated glyph shapes to the visualization of both scientific and information data. For information visualization, we have chosen an example of the visualization of "thematic" document similarities. Figure 2 shows a visualization of document similarities generated with the Telltale system.[12] The document corpus consists of 1883 articles from *The Wall Street Journal* from September and October 1989. Each glyph in figure 2 represents a document in the corpus, and the document's X, Y, and Z position, color and shape each represent the similarity of the document to one of the 5 themes.

Document similarity to *gold prices*, the *foreign exchange rate of the U.S. dollar*, and *federal reserve* are respectively mapped to the X, Y, and Z axes. The Y axis is visually indicated in figure 2 by the vertical line, with the X axis going to the right and the Z axis going to the left. The bulk of the documents have very low similarity to all of these 3 themes, so their glyphs are clustered near the origin at the bottom center.

The documents outside this cluster exhibit two spatial patterns: a cluster of 9 documents to the bottom right and a vertical branch on the left. The right cluster indicates the small number of documents in the corpus that discuss both *gold prices* and the *foreign exchange rate of the U.S. dollar*. The vertical branch depicts a larger collection of documents that discuss both *foreign exchange rate of the U.S. dollar* and the *federal reserve*.

A fourth attribute, similarity to *stock prices*, is inversely mapped to both superquadric exponents of the glyph shape, with highest similarity creating cuboids, then spheres, diamonds, and stars (lowest). Referring to the square array of sample glyphs in figure 1, the similarity to *stock prices* maps to glyphs on the diagonal from the upper left to the lower right of figure 1, with upper left indicating high similarity, and lower right indicating low similarity.

In figure 2 the larger, rounder shapes along the vertical branch exhibit some significant relationship to *stock prices* while the more numerous star-shaped glyphs do not. Clearly the vertical branch contains articles relating *foreign exchange*, *federal reserve* and *stock prices*.

Glyph color is mapped inversely to similarity to *Manuel Noriega*. Most of the documents fall in the turquoise and purple range, indicating no significant relationship. However, the documents in the orange, red, and yellow-green range represent documents with a significant relationship to *Manuel Noriega*. Many of these documents mention the effect of the coup attempt against *Manuel Noriega* and its effect on *foreign exchange rate of the U.S. dollar* (vertical axis). The fact that these orange, red, and yellow-green documents are not in either of the branches indicates that these articles did not relate heavily to either *federal reserve* or *gold prices*, and their star shape indicates no relationship with *stock prices*.

4.4. Scientific Visualization Results

We have used this system to examine several scientific visualization data sets. Figure 3 shows the visualization of a magnetohydrodynamics simulation of the solar wind in the distant heliosphere (20 times the distance of the Earth to the Sun). The simulation data is a $64 \times 64 \times 64$ grid containing the vector vorticity and velocity for the simulation.

Opacity is used to represent vorticity in the j direction, so that the 6 vortex tubes (only 4 are visible) represent zones in space where this vorticity is somewhat larger than zero. Glyph shape is based inversely on the velocity in the j direction. Positive velocities are displayed as larger, rounder to cuboid shapes and negative velocities are displayed as spiky, star-like shapes. Zero velocity is represented by the diamond shape. The overall columnar pattern of the data is not disturbed by the introduction of the shape mapping, but the velocity variation can still be seen as we traverse the lengths of the tubes. In this case, values close to zero in terms of j vorticity (still fluid) have been masked out.

Figure 4 is a visualization of the same magnetohydrodynamics data but with the opacity, color and glyph shape all mapped to the j component of vorticity. Negative vorticity components produce concave shapes (blue stars), while positive values produce convex shapes (orange cuboids and ellipsoids). Using this data mapping clearly shows three tubes with negative j vorticity and three tubes with positive j vorticity.

5. CONCLUSIONS

We have developed a new technique for intuitive, comprehensible creation of glyph shapes. This technique is based on procedurally generated superquadric functions and increases the number of dimensions of data that can be comprehensibly visualized in a glyph-based visualization system. These shapes allow the intuitive understanding of data variation among glyphs while preserving the global data patterns. We have shown the value of these techniques for both multi-dimensional information and scientific visualization.

REFERENCES

1. F. J. Post, T. van Walsum, F. H. Post, and D. Silver, "Iconic techniques for feature visualization," in *Proceedings Visualization '95*, pp. 288–295, October 1995.
2. W. Ribarsky, E. Ayers, J. Eble, and S. Mukherja, "Glyphmaker: creating customized visualizations of complex data," *IEEE Computer* 27, pp. 57–64, July 1994.
3. D. Ebert, C. Shaw, A. Zwa, and C. Starr, "Two-handed interactive stereoscopic visualization," *Proceedings IEEE Visualization '96* , October 1996.
4. W. S. Cleveland, *The Elements of Graphing Data*, Wadsworth Advanced Books and Software, Monterey, Ca., 1985.

5. J. Bertin, “Semiology of graphics,” 1983.
6. J. D. Foley and C. F. McMath, “Dynamic process visualization,” *IEEE Computer Graphics and Applications* 6, pp. 16–25, Mar. 1986.
7. A. Barr, “Superquadrics and angle-preserving transformations,” *IEEE Computer Graphics and Applications* 1(1), pp. 11–23, 1981.
8. D. S. Ebert, “Advanced geometric modeling,” in *The Computer Science and Engineering Handbook*, A. Tucker Jr., ed., ch. 56, CRC Press, 1997.
9. A. J. Parker, C. Christou, B. G. Cumming, E. B. Johnston, M. J. Hawken, and A. Zisserman, “The analysis of 3D shape: Psychophysical principles and neural mechanisms,” in *Understanding Vision*, G. W. Humphreys, ed., ch. 8, Blackwell, 1992.
10. H. Senay and E. Ignatius, “A knowledge-based system for visualization design,” *IEEE Computer Graphics and Applications* 14, pp. 36–47, November 1994.
11. H. Senay and E. Ignatius, “Rules and principles of scientific data visualization,” *ACM SIGGRAPH HyperVis Project*, *http://ironduke.cs.gsu.edu/classes/hypervis/percept/visrules.htm* , 1996.
12. C. E. Pearce and C. Nicholas, “TELLTALE: Experiments in a Dynamic Hypertext Environment for Degraded and Multilingual Data,” *Journal of the American Society for Information Science (JASIS)* 47, April 1996.

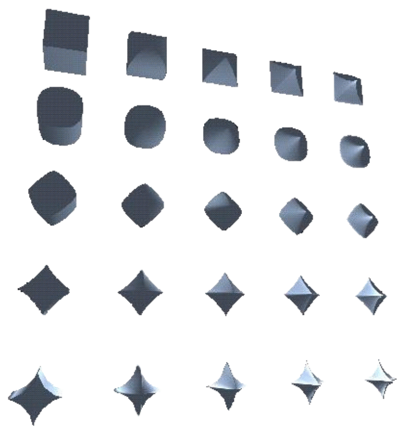

Figure 1. Example superquadric shapes created by varyin each exponent from 0 to 4.

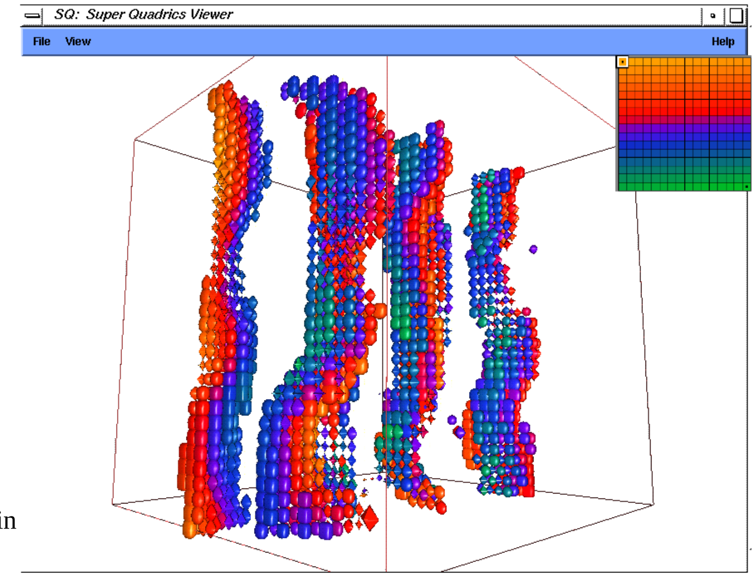

Figure 3. Visualization of a magnetohydrodynamics simulation of the solar wind in the distant heliosphere showing both velocity components and vorticity components of 6 vortex tubes.

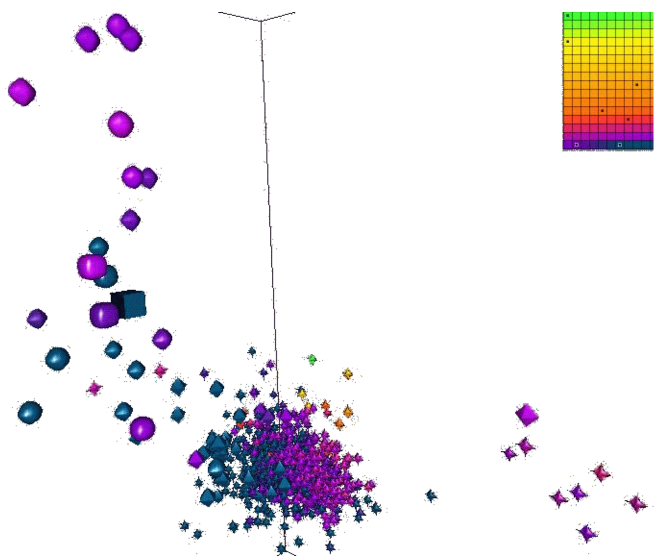

Figure 2. Three-dimensional visualization of 1833 documents' relationship to gold prices, foreign exchange, the federal reserve, stock prices, and Manuel Noriega.

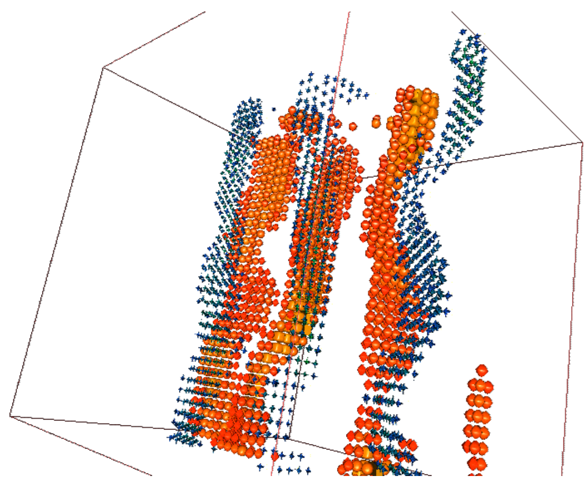

Figure 4. Visualization of a magnetohydrodynamics simulation of the solar wind in the distant heliosphere displaying 3 vortex tubes with positive j vorticity (cuboids and ellipsoids) and 3 vortex tubes with negative j vorticity (stars).