Using Shape to Visualize Multivariate Data

Christopher D. Shaw, James A. Hall, Christine Blahut, U. of Regina *

David S. Ebert, U. of Maryland Baltimore County †

D. Aaron Roberts, NASA Goddard ‡

Abstract

This paper describes our recent findings in the area of using glyph shape to display one or two data dimensions in the visualization of 3D scalar and vector fields. In our glyph-based visualization system, each glyph represents a data point in 3D space. Visual attributes such as size, orientation, color and transparency can be mapped to data dimensions in the 3D space. We are exploring the use of glyph shape as a display dimension, using superquadric superellipses as a means of supplying a parameterizable shape. A basic factor in effectively using shape for quantitative visualization is determining how many (and which) superellipse shapes people can distinguish. Since the superquadric shape's parameter set is not perceptually linear, we conducted a user study to which shapes people can generally distinguish. The findings show that with large superellipses, about 22 separate shapes can be distinguished on average. These results provide the foundation for exploring how effective superellipses may be in quantitative shape visualizaton.

1 Introduction

Our goal is the visualization of complex informations spaces in the realm of traditional scientific visualization, and in the emerging field of information visualization. Our emphasis has been on glyph-based visualization, where each data item is represented by a small visual symbol or *glyph* [8, 9, 5]. Glyphs have been successfully used to indicate flow, and they allow multiple data values to be encoded in their visual parameters. Glyph rendering is an extension to the use of glyphs and icons in numerous fields, including cartography, logic, and statistics.

In this paper, We explore the perceptibility of glyph shapes generated by superquardics [1] as a display mechanism. We have recently developed new techniques for automatic glyph shape generation that allow perceptualization of data though shape variation. Cleveland [3] 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. The next opportunity for encoding values 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. 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 [2]. We have chosen the procedural generation of glyph shapes using superquadrics [1] because superquadrics offer the required shape interpolation mechanism.

2 Procedural Shape 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 [4]. Superquadrics are a natural choice to satisfy this goal. Superquadrics are extensions of quadric surfaces where the trigonometric terms are each raised to exponents. Our implementation uses superellipses, because certain parameters will produce *box*, *sphere*, and *diamond* (regular octahedron) shapes.

For superellipsoids, the trigonometric terms are assigned exponents as follows:

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

These exponents allow continuous control over the convexity or concavity of the shape. Figure 1 shows a sample of the shape vocabulary. The exponents are (0,0) at the top left, and (4,4) at the bottom right.

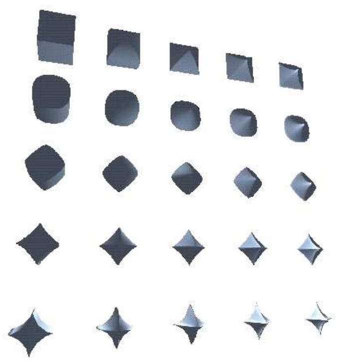

Figure 1: Example superquadric shapes created by varying each exponent from 0 to 4.

Glyph shape is a valuable visualization component because of the human visual system's pre-attentive ability to discern shape.

*Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada S4S 0A2, Phone: (306) 585-4071, email: cdshaw@cs.uregina.ca, hallj@cs.uregina.ca

†Computer Science and Electrical Engineering Department, University of Maryland Baltimore County, 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

Shapes can be distinguished at the pre-attentive stage [7] using curvature information of the 2D silhouette contour and, for 3D objects, curvature information from surface shading [6]. 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.

3 Experimental Evaluation

We conducted a user test to find the set of superellipsoid parameters that most people would be able to distinguish. The central question being addressed is: can a JND (Just-Noticeable Difference) scale be established so the user can pre-attentively identify a meaningful difference among superellipsoids, or a pattern of superellipsoids. The experimental approach was based on the Weber-Fechner theory of Just-Noticeable Differences. This theory, established in the early 1900's, focuses on the absolute threshold: how much a stimulus must change in order for the observer to sense a difference. The difference detected is the difference threshold. JND scales have been established for many factors, such as intensity of tone (1/10), and pressure (1/7). We propose to establish an incremented scale for superellipsoids that are noticeable difference from one another.

The experimental setup was a computerized JND task in which a target superellipsoid and an alterable superellipsoid were presented to the participant (figure 2). The participant was instructed to match the alterable figure to the target figure in shape. The participants used the mouse to control a slider scale directly below both figures. They dragged the slider either left or right along the scale to modify the shape and size of the alterable superellipsoid.

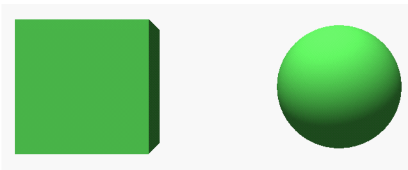

Figure 2: Example superquadric shapes created by varying each exponent from 0 to 4.

The computerized JND task consisted of three blocks of tests. The blocks were divided into three sub-blocks, each containing twenty trials. The first sub-block and the third sub-block, horizontal layout or vertical layouts, were counterbalanced. The horizontal layout (figure 2) contained two objects side by side, and the left object was the target. The instructions for the vertical layout were similar, with the target superellipsoid above the alterable figure. Each individual matching task had to be performed within a 10-second time limit; all times were recorded. Consequentially, if the participant had not completed the match within 10 seconds, a pop-up alarm box stating that time had elapsed would appear, trial was recorded but not statistically analyzed.

Fifteen women and sixteen men, between the ages of 18 and 48 years old, participated in the study. All participants were post-secondary university students. The average age of the males who participated in the experiment was 24.13 years (SD=4.79); the average age of the females was 24.40 years (SD= 7.07).

Target	Mean Response	Std Dev
0.0	0.0015	0.0061
0.3	0.3203	0.0934
0.5	0.5304	0.0972
0.7	0.6620	0.0738
0.8	0.7780	0.1064
1.0	1.0089	0.0225
1.1	1.1432	0.0570
1.2	1.2892	0.1258
1.6	1.6162	0.1025
1.8	1.8330	0.1166
2.0	2.0565	0.1101
2.3	2.3122	0.1792
2.5	2.5673	0.1664
2.7	2.7471	0.1780
2.9	2.9329	0.1745
3.0	3.0815	0.1843
3.4	3.3760	0.1839
3.7	3.5413	0.1993
3.9	3.7443	0.1662
4.0	3.8064	0.1919

Table 1: Superquadric target parameters with the mean and standard deviation of the response.

3.1 Targets

Twenty target objects were presented in random order, and the participant was asked to match the target with the adjustable superellipsoid. We used the same 20 targets across all trials so that statistical analysis could be performed per target.

To simplify the search space, we used the superellipsoids that were symmetric about all 3 axes. This yields the shapes along the diagonal from the top left to the bottom right in figure 1. This is accomplished by setting the two superellipsoid exponents to the same value at all times. That is, $\epsilon 1 = \epsilon 2$ for all the superellipsoids in this experiment.

To bootstrap the process of selecting a set of targets for each participant to match, we first chose 20 candidate parameters (values of $\epsilon 1 == \epsilon 2$) from a simple linear ramp: $0.0, 0.2, 0.4, \ldots 4.0$. We then performed some trial runs of the experiment using these parameters, and adjusted some parameter values up or down by one or two tenths to reflect our impression that some parameter ranges did not seem to change the shape much.

3.2 Results

Table 1 shows the experimental parameter plus the experimental outcome.

The graph in figure 3 shows the standard deviations that were measured in user responses at each target. The interesting points to note are that the standard deviation at $\epsilon 1 = 0.0$ is very small. This is because the left limit of the adjustment slider was $\epsilon 1 = 0.0$, giving participants a strong indication that they had reached the correct parameter. Also, the shape at $\epsilon 1 = 0.0$ is a box, which is readily recognizeable.

The standard deviation at $\epsilon 1 = 1.0$ is also very small. This value of $\epsilon 1$ yields a sphere, so it seems that participants were quite able to discriminate a sphere from the surrounding "not-quite sphere" shapes.

Aside from $\epsilon 1 = 0.0$ and $\epsilon 1 = 1.0$, the values of $\epsilon 1$ between 0 and 2 tended to yield a standard deviation around 0.1. Beyond $\epsilon 1 = 2.0$, the standard deviation is around 0.18, with a rising trend. Interestingly, although the right limit of the slider is at $\epsilon 1 = 4.0$, there is still a large standard deviation, indicating that participants

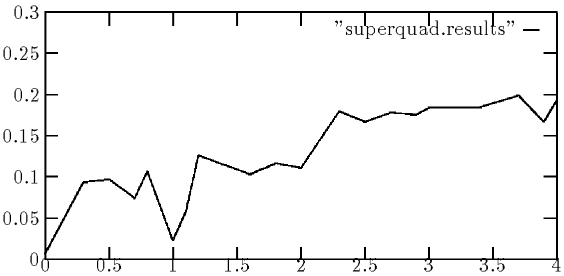

Figure 3: Plot of Standard Deviations. The horizontal axis is the superquadric parameter value, and the vertical axis shows the experimentally measured standard deviation.

genuinely did not know that the target shape was equal to the shape at the slider limit.

3.3 Estimating Number of JND Shapes

The goal of this study is to determine how many shapes are discriminable, and what the parameter values are of those shapes. This set of shapes would then form a table of shapes that could be selected by a data parameter in a visualzation system. For example, if there were 10 distinct shapes, then given a data parameter in the range $[0..1]$, then values $[0..0.1]$ would select shape 0, values $[0.1..0.2]$ would select shape 1, and so on.

To determine this set of shape parameters, we computed a recursive linear estimation of paramater values. If we repeated the experiment with this new set of target values, participants would never be off in their response by more than 1 shape step. The data indicate that the participant responses are normally distributed about the means. Thus, setting the next estimated $\epsilon 1$ value to be 2 standard deviations away should yield a situation in which 68% of responses for each target should be correct, and almost all responses should be off by only one target.

Starting at one endpoint of the parameter space (say $\epsilon 1 = 4.0$), we then computed the next Just-Noticeably Different shape by subtracting $2 \times StdDev$. We in fact need to compute the standard deviation as it exists around the next parameter, since the standard deviation is not equal for all parameters:

$$\epsilon_{dest} = StdDev_{dest} + StdDev_{source} + \epsilon_{source} \qquad (2)$$

Since $StdDev_{dest}$ is not yet known, we used an altered form of Newton's root-finding method to compute it:

$$\epsilon'_{dest} = StdDev_{source} + StdDev_{source} + \epsilon_{source} \qquad (3)$$

ϵ'_{dest} is the first guess at the destination ϵ_{dest}. There will be a pair of target ϵ values named ϵ'_{left} and ϵ'_{right} that bracket ϵ'_{dest}, and these will have their associated standard deviations. The value of ϵ'_{dest} is used to linearly blend between the two standard deviations:

$$\alpha = \frac{\epsilon'_{dest} - \epsilon'_{left}}{\epsilon'_{right} - \epsilon'_{left}} \qquad (4)$$

$$StdDev_{dest} = (1 - \alpha) \times StdDev_{left} + \alpha \times StdDev_{right} \qquad (5)$$

Finally, the estimated $StdDev_{dest}$ is used to compute ϵ_{dest}, as shown in equation 2.

The new ϵ_{dest} is then put on the role of ϵ_{source}, and the preceding process is repeated to generate the desired list of perceptually-linear superellipsoid shapes. In situations where standard deviations don't change much, this procedure is reasonably accurate. With changing standard deviations, repeated iterations of equations 3 through 5 could make the estimate more accurate, but since the basis of this estimation procedure is statistically variable data, we did not do this.

Another issue is which endpoint to start at. We computed two parameter tables; one starting at $\epsilon 1 = 0.0$ and working up, and one starting at $\epsilon 1 = 4.0$ working down. Table 2 shows these results.

Starting at $\epsilon 1 = 0.0$	Starting at $\epsilon 1 = 4.0$	Mean
0	0.010	0.005
0.016	0.030	0.023
0.042	0.063	0.052
0.086	0.117	0.102
0.161	0.206	0.184
0.289	0.355	0.322
0.466	0.541	0.504
0.641	0.727	0.684
0.825	0.881	0.853
0.966	0.975	0.971
1.037	1.042	1.039
1.133	1.177	1.155
1.318	1.393	1.355
1.545	1.610	1.577
1.763	1.828	1.795
1.987	2.054	2.020
2.255	2.341	2.298
2.928	3.056	2.639
2.583	2.695	2.992
3.286	3.425	3.356
3.667	3.805	3.736
4.072	4.000	4.036

Table 2: Estimated Superquadric parameters derived from repeatedly adding 2 standard deviations to the current parameter value to generate the new one. One of these standard deviations is computed using a linear interpolation. The third column shows the mean of these two estimates.

4 Conclusions

The results show that there are 22 distinct superellipsoid shapes, with 10 in the range $0.0 < \epsilon 1 < 1.0$, 5 in the range $1.0 < \epsilon 1 < 2.0$, and 7 in the range $2.0 < \epsilon 1 < 4.0$.

We are in progress towards using these derived parameters in an experiment to test for preattentive shape finding. 35 identical distractor superellipsoids will be presented in a perspective view with 1 different superellipsoid. The participant will select the different superellipsoid and participant response time will be collected.

5 Acknowledgements

This project was supported by NSERC Canada under grant OGP0183941, and by NASA under grant NAG 5-2893.

References

[1] A. Barr. Superquadrics and angle-preserving transformations. *IEEE Computer Graphics and Applications*, 1(1):11–23, 1981.

[2] J. Bertin. Semiology of graphics, 1983.

[3] William S. Cleveland. *The Elements of Graphing Data*. Wadsworth Advanced Books and Software, Monterey, Ca., 1985.

[4] David S. Ebert. Advanced geometric modeling. In Allen Tucker Jr., editor, *The Computer Science and Engineering Handbook*, chapter 56. CRC Press, 1997.

[5] J. D. Foley and C. F. McMath. Dynamic process visualization. *IEEE Computer Graphics and Applications*, 6(3):16–25, March 1986.

[6] Victoria Interrante, Penny Rheingans, James Ferwerda, Rich Gossweiler, and Toms Filsinger. Principles of Visual Perception and its Applications in Computer Graphics. In *SIGGRAPH 97 Course Notes, No. 33*. ACM SIGGRAPH, August 1997.

[7] Andrew J Parker, Chris Christou, Bruce G Cumming, Elizabeth B Johnston, Michael J Hawken, and Andrew Zisserman. The Analysis of 3D Shape: Psychophysical Principles and Neural Mechanisms. In Glyn W Humphreys, editor, *Understanding Vision*, chapter 8. Blackwell, 1992.

[8] Frank J. Post, Theo van Walsum, Frits H. Post, and Deborah Silver. Iconic techniques for feature visualization. In *Proceedings Visualization '95*, pages 288–295, October 1995.

[9] W. Ribarsky, E. Ayers, J. Eble, and S. Mukherja. Glyphmaker: creating customized visualizations of complex data. *IEEE Computer*, 27(7):57–64, July 1994.