

Simple 3D Glyphs for Spatial Multivariate Data

Camilla Forsell¹
Dept. of Information Science
Uppsala University
Uppsala, Sweden,

Stefan Seipel²
Dept. of Mathematics, Natural- and
Computer Science
University of Gävle
Gävle, Sweden

Mats Lind³
Dept. of Information Science
Uppsala University
Uppsala, Sweden

ABSTRACT

We present an effort to evaluate the possible utility of a new type of 3D glyphs intended for visualizations of multivariate spatial data. They are based on results from vision research suggesting that our perception of metric 3D structure is distorted and imprecise relative to the actual scene before us (e.g., [1]); only a class of qualitative properties of the scene is perceived with accuracy. These properties are best characterized as being invariant over affine but not Euclidean transformations. They are related, but not identical to, the non-accidental properties (NAPs) described by Lowe [2] on which the notion of geons is based [3]. A large number of possible 3D glyphs for the visualization of spatial data can be constructed using such properties. One group is based on the local sign of surface curvature. We investigated these properties in a visualization experiment. The results are promising and the implications for visualization are discussed.

CR Categories and Subject Descriptors: I. 3.6 [Computer Graphics]: Methodology and techniques: Interactive techniques.

Keywords: multidimensional visualization, perception, 3D glyphs.

1 INTRODUCTION

Information technology rapidly changes the ways in which information can be accessed and visualized. New sensor technology can provide us with near-real-time geo-spatial observations and measurements of the real world. This technology improves our ability to discover, access and integrate data and offer the promise of better and quicker decision making, as well as the potential to improve monitoring, detection and planning activities over a wide range of areas. The critical question is how to transform such raw data into visualizations that can be effectively and efficiently used as optimal support. In other words, to help users' discover the task-relevant information available and enable them to do it quickly and correctly.

Information obtained from a specific geographical location is often of a multivariate nature. Consider, for instance, meteorology. Here, one location produces data of several kinds

such as temperature, humidity, wind (speed and direction) and atmospheric pressure. If one also includes simulation programs (for possible course of events), the available amount of data to a decision maker often becomes immense. When visualizing multidimensional information, the human perceptual system sets a limit on the number of dimensions in a data set that can be meaningfully shown. Several ideas to extend the number of dimensions have been suggested over the years by, for instance: Chernoff [4] utilizing simple depictions of faces, Chambers, Cleveland, Kleiner and Tukey [5] using "star plots" and by Picket and Grinstein [6] using stick figures. Such visualizations could also be placed in a grid structure illustrating multivariate geo-spatial data. One further option is to explore the inherent nature of human perception and somehow use a 3D spatial layout for the data. However, as pointed out by Lind, Bingham and Forsell [1], 3D depictions, regardless of the type of display technique employed, must be used with discretion. A large set of 3D relationships are simply not perceived by humans although our subjective experience leads us to believe that our visual system delivers a complete and veridical representation of the scene before us. Research on human spatial vision shows that an observer's judgments of metric 3D shape and relationships contain large errors relative to the actual structure of the observed scene [1], [7] and [8]. Only structural properties in 3D left invariant by affine mappings are reliably perceived. Examples of such structural relationships are the relative sign of the curvature on a surface, the parallelism of lines or planes and relative distance intervals in parallel directions. Thus, the perceptual representation of 3D shape is likely to be based on such qualitative aspects of structure. All remaining aspects (of structure) are inherently ambiguous. From a visualization perspective, this implies that we can expect users to misinterpret or overlook some, even largely significant, relations in most 3D visualizations.

As we see it, these results do not imply that 3D visualizations should not be used, only that care must be taken when employing them. Because qualitative aspects of 3D structure are reliably perceived, conveying nominal or ordinal relationships by means of 3D structure should be an interesting option to investigate. In this vein we have looked at the possibility of constructing 3D glyphs that could depict ordinal data by using the sign of curvature of surface elements. In 2D a line segment could represent three ordinal levels by having a negative curvature, no curvature or a positive curvature. By employing this in two orthogonal directions, a small square surface patch in 3D could show three ordinal levels for two independent variables. The surface patch will then, of course, have one of nine different shapes (Figure 1). Using shape to depict ordinal data goes completely against the recommendations by McKinlay [10],

¹ Camilla.Forsell@dis.uu.se

² SSL@hig.se

³ Mats.Lind@dis.uu.se

something he considered should be avoided and being the worst possible choice. However, we still found merit in our idea for three reasons. First, the sign of curvature, and therefore also the resulting shapes, ought to be easily distinguishable from one another for theoretical reasons. Second, although each shape represents an integral impression of the two variables, they were, both to us and to other viewers of our sample visualizations, quite possible to visually investigate variable by variable. Perhaps this is because that only three levels for each variable are used. Third, the sign of curvature, at least to us, has an intuitive coupling to a three-level ordinal scale.

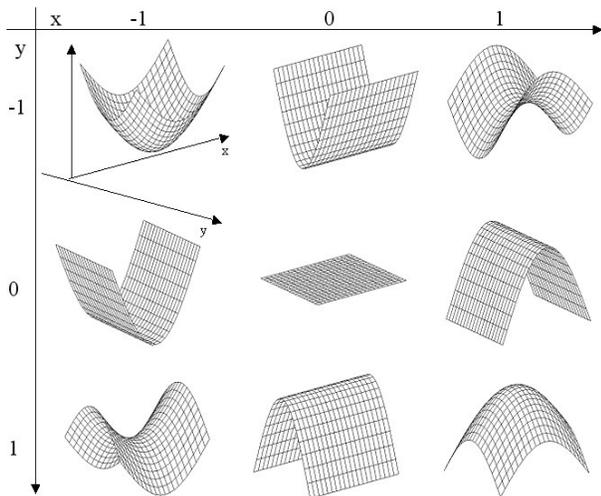


Figure 1. The nine shapes as defined by the sign of curvature.

To be useful in a real application, three levels are probably not enough, at least most of the time, and it is probably desirable to show more than two variables per grid cell. We therefore looked into possibilities of extending our idea to incorporate more levels and more variables per surface patch. To find more levels we looked at the possibility of using infinite curvature at the negative and positive side of curvature and in this manner produce five levels per variable. By using a discontinuity, an edge, this can be accomplished (Figure 2). A surface having an edge or not is also a qualitative property invariant over affine transformations, so it should be easy to perceive as well. Furthermore we think that these levels are easy to perceive as an ordered sequence, especially if a difference in height is superimposed. Height, naturally, is not invariant over affine transformations and thus not reliably perceived, but because it is used redundantly it will aid the perceiver, at least in some cases.



Figure 2. Five levels of potentially easily discriminable 2D curvatures.

To add one more variable we covered each surface with a color chosen from a sequence based on increasing luminance and differences in hue. A fourth, nominal, variable was also added by the use of different textures on the colored surface patches.

Before implementing these ideas in a larger scale, however, we wanted to empirically investigate their merit, doing it in an incremental fashion. Therefore, we started with the original and simpler three-level idea with an added three level color-coding to get a visualization with three ordinal variables each having three levels. Here we report the results of this first investigation.

2 TO EVALUATE THE EFFICIENCY OF VISUALIZATIONS OF SPATIAL MULTIVARIATE DATA

The effectiveness of visualizations will largely be determined by the number of fixations a viewer needs in order to extract the relevant information contained in the visualization. This, in turn, is largely determined by the number of repeated re-fixations between two objects in the scene that is needed to perceive the structural differences between them. Elsewhere [11], we have investigated the differences in orientation between two 2D squares needed for certain discrimination utilizing only one fixation per square. The results show that, even if humans are able to perceive orientation differences as small as 5 to 6 degrees under more or less unlimited viewing conditions [12], when restricted to one fixation per object the orientation difference needed is roughly 25 degrees. This means that if we use orientation of a square as a carrier of data in a visualization and only used two values represented by a square and a 45 degree rotated square (a diamond), a user of this visualization could scan through it using only one fixation per square. On the other hand, if several values were used resulting in orientation differences of only 10 degrees or so between the squares, a user would spend a considerably longer amount of time changing his or her fixation back and forth between the squares to determine if they were different. It could reasonably be argued that the longer time needed is compensated by the higher degree of accuracy obtained in the reading. However, you could also argue that, given our results, orientation differences are not suitable for more levels than two and that other forms of data representations should be considered in this case.

Undoubtedly, though, an important metric to consider when discussing the effectiveness of visualizations is the amount of re-fixations needed for a user to discover the relevant information. A similar approach was used by Lightner [13] when analyzing the effectiveness of visualizations. She used a simple model to calculate the minimum time needed for a user to successfully investigate a visualization:

$$T_{\min} = 200 + 230 \times (a - 1)$$

where a is the number of areas that needs to be fixated in the visualization to solve the task and T_{\min} is the minimum expected time, in msec, to solve this task. The logic behind the model is that a fixation usually lasts for around 200 msec and that an average saccade (movement of the eye to fixate a new point) takes about 30 msec. These figures are, of course, only approximations. A fixation can last up to 500 msec [14] and the time it takes for a saccade is dependent on the amount of rotation of the eyeball needed. The model can therefore be simplified further into:

$$T_{\min} = \frac{a}{4}; T_{\min} \text{ in seconds}$$

without any significant loss of resolution. However, even as such it can be extremely useful and we will use it to predict and interpret our empirical findings below.

3 EXPERIMENT

To evaluate the effectiveness of our proposed simple 3D glyphs we created a judgment situation to be performed as a free-scan visual search task. The data visualized were hypothetical meteorological data. There were three variables defined in each of 144 (12 by 12) grid cells symmetrically positioned over a square geographical area. The three variables were temperature, precipitation and wind speed. Each variable was categorized into three levels: “below average”, “average” and “above average”. Two of these variables, temperature and wind speed, were mapped onto the curvature of a surface patch for each grid cell. This surface patch would take on one of the nine possible shapes defined by the sign of curvature in the x and y directions (Figure 1). The third variable, precipitation, was mapped onto each surface patch by the use of color. A value of “below average” resulted in a dark blue surface patch; the value “average” resulted in a medium gray surface patch; and the value “above average” resulted in a bright red surface patch.

The 144 grid cells that could contain surface patches were clustered into 9 (3x3) main areas, each comprising 16 (4 x 4) individual grid cells (Figure 6) The observer’s task in all conditions was to identify which of these nine main areas contained at least four grid cells having surface patches of a specified kind called “target patches”. Only one main area met this condition per trial. All other individual grid cells had randomly chosen surface patches among the 27 possible types.

Forsell and Lind (in preparation) studied the relative discriminability of these surfaces. Most of the surfaces could be discriminated from the others during only one fixation per surface, even under extremely poor viewing conditions (top-view, monocular viewing, depth only conveyed by cast shadows). A slightly lower performance, however, was found with the two saddle shaped surfaces under these conditions. For this experiment, we therefore chose one of the “easy” surfaces and one of the “difficult” surfaces as target patches — the “top” and one of the “saddles”. In the terminology of the task these were described to the observers as “high risk for fire” and “high risk for contamination”. High risk for fire was defined as temperature above normal, precipitation below normal and wind above normal. In our visualization this was characterized as red saddles (a “difficult” surface) in a specific orientation. High risk for contamination was defined by all three variables having levels above normal. In our visualization red tops (a “simple” surface) characterized this condition.

The display used in the experiment, given our set up, subtended a visual angle of about 36 degrees squared. Because the display was horizontal and because the observers looked at it from one side, the visual angle subtended by the display was of course slightly different close to the observers as compared with further away from the observers. In that we used 12 by 12 individual grid cells, each grid cell subtended a mean visual angle of about 3 by 3 degrees. The visual acuity of the human visual system is very nonuniformly distributed over the visual field. Already at 10 degrees from the fovea we can only resolve 1/5 the detail compared with the fovea [15 p. 59]. A reasonable approximation is therefore that our surfaces patches can be discriminated from one another with certainty within a solid angle in the visual field having a diameter of somewhere between 5 and 10 degrees. A separate experiment would be needed to determine a more exact value. This means that somewhere between 1 and 4 grid cells can

be searched per fixation. Because our task was such that the search could be terminated once an area with four target patches was found, and random placement was used for target areas, the mean number of surface patches needed to be searched was $144/2 = 72$. Using these numbers in our simple model, we can estimate that a minimal search time for finding the target area in the display is between $[(72/4) / 4] = 4.5$ sec and $[(72/1)/4] = 18.5$ sec. If the time needed to find the target area is longer than this range it is a strong indication that a non-negligible amount of re-fixations is needed. Our prediction for the experiment, given our theoretical considerations, is however that the time we obtain for our observers to solve a trial will be within the 4.5 to 18.5 sec range, indicating that only few re-fixations are necessary.

A common method of depicting grid cell values in geographical information systems is to color grid cells according to some color scheme. If there is more than one variable defined over the geographical area, these systems often force you to look at them one at a time using the same color scheme. As a simple control condition, we therefore introduced a visualization that contained three layers over our square display surface. Each layer consisted of the same 12 by 12 grid cells as in the 3D condition but here no surface patches were added. Instead, each layer was assigned to a variable and the color scheme used for one variable in the 3D condition was employed for each of the layers, i.e., for all of the three variables, which could then be viewed one at a time by switching layers. To improve the usability of this condition, three specially designated keys were used so that each layer could be addressed with the press of a single button. Figure 3 shows a trial in the 3D condition while Figure 4 depicts the layers of the 2D condition, here shown adjacent to one another.

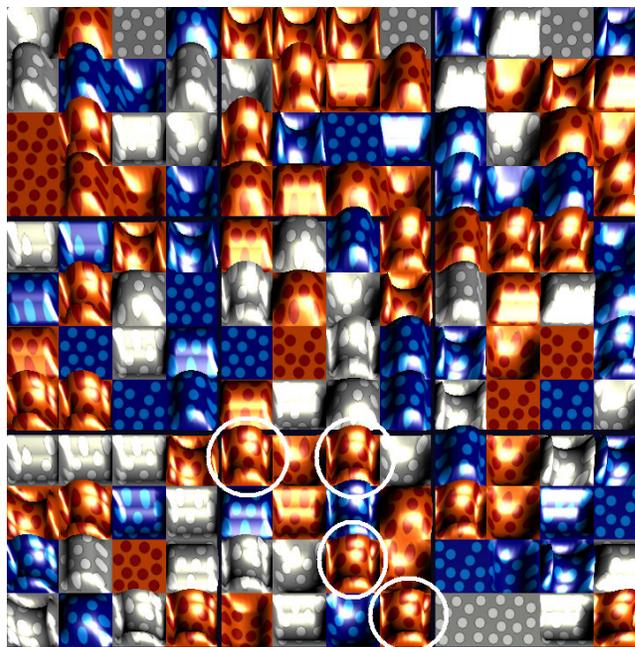


Figure 3. A screen shot from the 3D condition. Target state for search is “high risk for fire”. Here it is characterized by four red saddles in one of two possible orientations within the same area. The target is located in the middle area of the bottom row as illustrated.

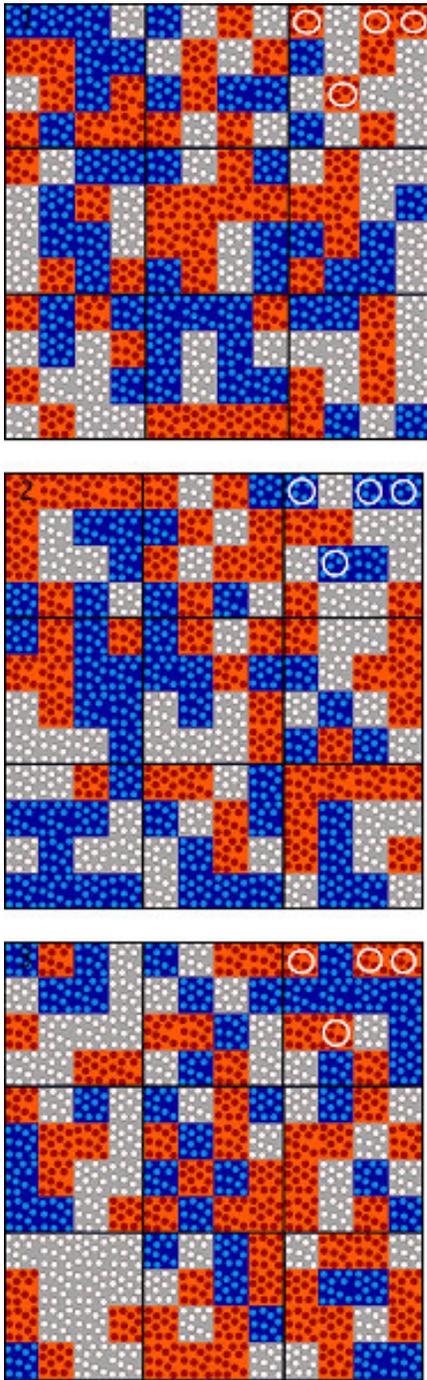


Figure 4. Screen shots from the 2D condition. The target state for search is, as in Figure 3, “high risk for fire”. The target area is located in the top right corner. It contains four grid cells in corresponding locations in the three separate layers, with the variables colored as defined by the target state: red for temperature (top display), blue for precipitation (middle display) and red for wind speed (lower display).

3.1 Method

3.1.1 Stimulus

As described above, each stimulus display was comprised of a 12 by 12 matrix of grid cells creating a total of 144 grid positions with a square size of 0.8 by 0.8 meters subtending a visual angle

of approximately 36 degrees. The matrix was divided into 9 main areas, each with 16 grid positions. Each grid position was assigned a set of values that defined its shape and color. In one of the nine areas four grid cells contained a target surface patch. All other grid cells contained randomly sampled surface patches from the 27 possible combinations (3 x 3 x 3). There was one parameter controlling the random sampling; no other main area, of the nine, contained more than three target surface patches.

The images were rendered using Open GL. A polka dot texture with two shades of the designated color was added to each surface patch to provide one further cue to 3D structure.

3.1.2 Apparatus and viewing conditions

The physical display system used was an integrated retro-projection display system in which the projection screen is oriented horizontally. This approach provides a viewing metaphor of a virtual world that is perceived as a bird’s eye view. The display environment has a square screen area of 0.8 by 0.8 meters. It can simultaneously project the content of eight co-located and interactive views and thus provide four independent observers with individual stereoscopic imagery. For our purposes, we used only one of these views (two projectors). The visible pixel resolution of each projection image on screen was 768 by 768 pixels. The separation of stereo image pairs was accomplished by using polarizing filters in front of the image projectors. The display system was driven by a small cluster of four commercial off-the-shelf computers interconnected with a Giga-Ethernet local area network. The computers were equipped with Nvidia QuadroFX graphics cards. For the purpose of head position tracking a magnetic tracking system, connected to a separate computer, was used (the Ascension Flock of Birds).

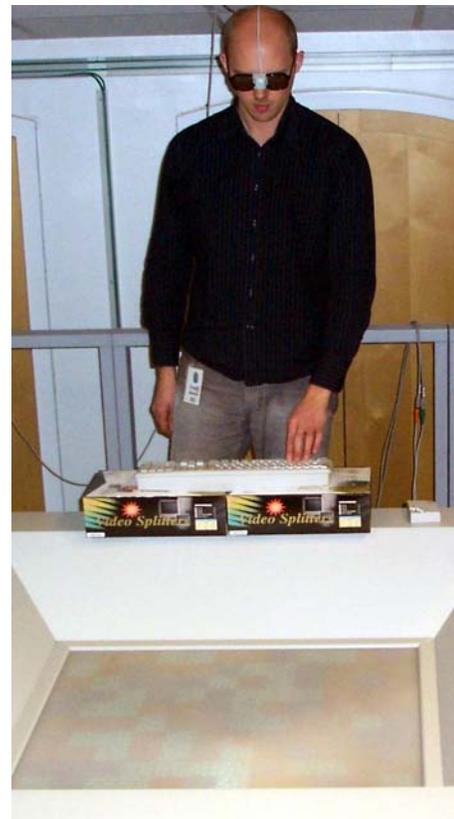


Figure 5. The experimental set up.

The observers were standing in front of the screen, which meant they were viewing the stimulus from an upright position (Figure 5). They wore polarizing glasses with a position sensor attached. Movement along the side of the screen was allowed. Thus motion parallax was introduced to further enhance the stereoscopic illusion. A keyboard used for responses was placed in front of the observer. Dim lighting was used in the room.

3.1.3 Experimental design

The study was designed as a five-factor mixed factorial design with three within-subject factors: viewing condition (2D/3D), target-glyphs (top/saddle) and block of trials. The between-subject factors were two, order of presentation of viewing condition and order of presentation of target-glyphs. The experiment was performed over four separate sessions, each consisting of three blocks of three trials. In order to prevent effects of any use of search strategies, these blocks were balanced with respect to target placement. The presentation order of the four separate sessions was counterbalanced using a Latin-square procedure. Half of the observers started with one of the two 3D conditions while the other half started with one of the two 2D conditions. This design yielded a total of 36 trials per observer.

3.1.4 Procedure

Before the start of the experiment the observers were shown printed, black and white wire-frame images representing the nine possible glyphs. They were extensively instructed about the geometric definition of the different types of curvature and orientation depicted. The experimenter explained the task, demonstrated the response apparatus to the observer and then the experiment commenced. The observers were told that their response times would be measured and that they should therefore respond as quickly as possible while maintaining high accuracy. Before each of the four separate sessions the observers were given written instructions about the task and completed a short block of practice trials. Response time was measured from the onset of the stimulus to the response. The sequence of trials was self-paced, and the observers controlled the screen by pressing the space bar when they wished a trial to begin. A blank screen was immediately presented after a response and at the start of each session. Directly after the space bar was pressed the search image appeared.

In the 2D sessions with three search images to be judged and compared within each trial the observer alternated between these images with keys dedicated to display each one at a time. The three images were labeled in their top left corners as an aid for the observer to keep track: 1 for temperature, 2 for precipitation and 3 for wind speed respectively. The alphabetical keys Z, X and C were used for navigation and relabeled with 1, 2 and 3 respectively. The observer's task in any individual trial was to search the image (3D) or the images (2D) and find the correct target area as specified by the target state described earlier. Stimuli were displayed until a response was given. This completed a trial and a blank screen reappeared. The response keys used were the numbers 1 to 9 on the numerical keyboard and chosen such that each key corresponded to one area on the stimulus display. Each area corresponded to the numerical keyboard with identical numbering, starting from the bottom of the display to the top and from left to right, creating a logical mapping between stimulus and response apparatus (Figure 6). To initiate a new trial the observer pressed the space bar. Errors and reaction times for each trial were recorded. No feedback was provided

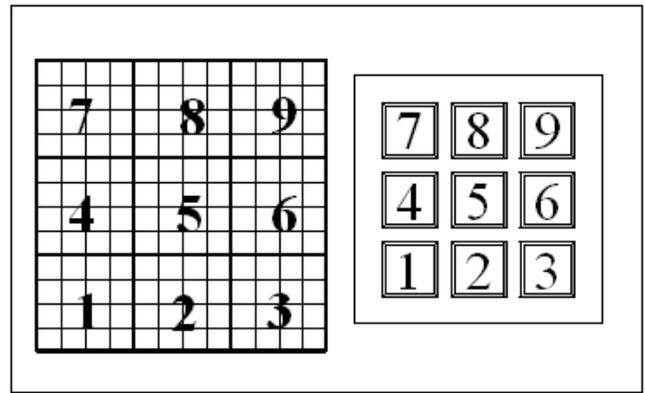


Figure 6. The numerical keyboard was used as the response apparatus and the observers were instructed to consider the areas in a trial display as having identical numbering. Hence, when the target of a search was located in area nine the observer pressed key "9", etc.

3.1.5 Observers

Twelve male observers, aged between 24 and 43 years old, took part in the experiment. All but one were students or staff of the Swedish National Defense College (SNDC). They had no prior knowledge of the purpose of the experiment or the specific hypotheses tested. All observers had normal or corrected to normal vision. They received a small compensation for taking part in the experiment.

3.1.6 Results

First, we analyzed the error data in the 3D condition. Errors were scarce; the group mean value was as small as 0.42 errors per 18 trials and the largest amount of errors for any observer was 2 out of the 18 trials, suggesting that the search times are good representatives of performance. The group mean value for the search times in the 3D condition was 9.12 sec, which is well within the range we specified for efficient performance.

Next, we turned to the comparison with the 2D condition. To begin with we compared the error rates. Because this type of data cannot be considered ideal for parametric testing, an exact test was used. First, the number of errors for the whole 18 trials in the 2D condition was determined for each observer; then the corresponding numbers for the 3D condition were compared. As stated above, the group mean value for the 3D condition was 0.42 errors per 18 trials. For the 2D condition this value was 1.75. The exact test, i.e. going through all 4096 permutations of values between the two conditions and for all observers, revealed that the probability ("two-tailed") of finding this large, or larger, absolute difference in the group mean value by chance is less than 0.024. We therefore conclude that, for our observers, the 2D condition produces more errors.

Because reaction time data are typically not normally distributed, we employed a logarithmic transformation of these data before further statistical testing. We also collapsed each observer's data over each set of three trials because such a set is less contaminated by possible response preferences in that it contains a balancing of the occurrence of the correct position within the grid. A mixed ANOVA was carried out using a decision criterion of 0.05 and with the order conditions as between-subjects factors and 2D/3D, target type and block as within-subjects factors. No main effect of the ordering factors was

found. The 3D times were significantly faster than the 2D times ($F(1,8)=787.7$, $p < 0.0001$). No significant difference was found between the two targets shapes ($F(1,8)=1.499$, $p < 0.26$), nor any interaction between target shapes and the 2D/3D condition ($F(1,8)=1.618$, $p < 0.24$). There was, however, a significant effect of block ($F(2,16)=5.398$, $p < 0.02$). The main finding (i.e. the main effect of the 2D/3D condition) is summarized in Figure 7.

It is noteworthy that the group mean search time for a correct response in the 2D condition (57.7 sec) was more than six times longer than in the 3D condition (9.1 sec). We do not suggest that this is a fair comparison. A more logical comparison would be to use a 2D glyph intended for multivariate data. However, in our study it served the purpose of comparing our visualization to the one actually used at the SNDC where the study was conducted.

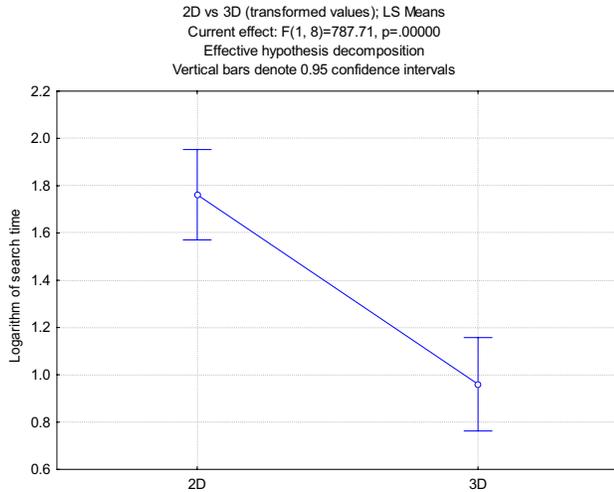


Figure 7. The main effect of the 2D/3D condition. Please note the search time values are logarithmic. The group mean value for the 2D condition is 57.7 sec and the group mean value for the 3D condition is 9.1 sec.

4 DISCUSSION AND CONCLUSION

The primary goal of this study was to determine whether observers would make efficient use of the 3D structural information we made available, i.e. to discriminate between the 3D glyphs and identify the correct main target area among the alternatives. The results of the experiment suggest that our observers were indeed able to do make this discrimination. In this richer visual environment no difference was found between the “top” and the “saddle”, something we found earlier under much more impoverished conditions. Our conclusion is that the results clearly demonstrate that affine 3D properties can be successfully used in some 3D visualizations of multivariate data and that further examinations of the perception of 3D glyphs and their potential use as elements for visualization are worthwhile.

It should be noted that most of the non-accidental properties (NAP:s) identified by Lowe [2], in a quest for view-point invariant cues to 3D shape can be expressed as structural properties invariant over affine transformations. Therefore, our proposed shapes are relatives of the geons proposed by Biederman [e.g., 3] to be the building blocks of mental representations of objects. However, other structural relationships than NAPs could also be utilized. For instance, the midpoint between two points on a straight line (bisection) is a property that is invariant over affine

transformations. We will continue to investigate the possible use of such structural properties and their role in creating efficient 3D visualizations and hope to reach a point soon where recommendations for use outside of the lab can be made.

ACKNOWLEDGEMENTS

This research was in part sponsored by project AQUA at the Swedish National Defense College in Stockholm, Sweden.

REFERENCES

- [1] Mats Lind, Geoffrey P. Bingham and Camilla Forsell. Metric 3D structure in visualizations. *Information Visualization*, 2:51-57, 2003.
- [2] David.G. Lowe Perceptual Organization and Visual recognition. Unpublished doctoral dissertation, Stanford University, Stanford, CA, 1984.
- [3] Irving Biederman. Recognition-by-components – A theory of image understanding. *Psychological Review*, 94 (2): 115-147, 1987.
- [4] Herman Chernoff. Using faces to represent points in k-dimensional space. *Journal of the American Statistical Association*, 68 :361-368, 1973.
- [5] John M. Chambers, William S. Cleveland, B Kleiner and Paul A. Tukey. *Graphical methods for data analysis*. Wadsworth, Belmont, CA, 1983.
- [6] Ronald M. Picket and George G. Grinstein. Iconographic displays for visualizing multidimensional data. In Proceedings of the 1988 IEEE Conference on Systems, Man and Cybernetics, volume 1, pages 415-519.
- [7] James T. Todd and J Farley. Norman. The visual perception of 3-D shape from multiple cues: Are observers capable of perceiving metric structure? *Perception & Psychophysics*, 65: 31-47, 2003.
- [8] James T. Todd, James S. Tittle and J Farley. Norman. Distortions of three-dimensional space in the perceptual analysis of motion and stereo. *Perception*, 24:75-86, 1995.
- [9] James T. Todd. The visual perception of 3D shape. *Trends in Cognitive Science*, 8 (3):115-121, 2004.
- [10] Jock D. Mackinlay. Automating the design of graphical presentations of relational information. In Stuart K. Card, Jock D. Macinlay and Ben Schneiderman (Eds), *Information Visualization: Using vision to think*, Morgan Kaufman Publishers: San Francisco, CA, 1999.
- [11] Mats Lind, Camilla Forsell and Alexander Allard. Effective visualizations for large displays – The role of transsaccadic memory. *Proceedings of the VIIP 2003 conference*, Benalmadena, Spain, September 8-10, 2003
- [12] Terry Caelli and Paul Bevan,. Probing the spatial frequency spectrum for orientation sensitivity in stochastic textures. *Vision Research*, 23: 39-45, 1983.
- [13] Nancy J. Lightner. Model testing of users’ comprehension in graphical animation: The effect of speed and focus areas. *International Journal of Human-Computer Interaction*, 13(1):53-73, 2001.
- [14] Julian Hochberg. Perception as purposeful inquiry. *Behavior and Brain Sciences*, 22: 619-620, 1999.
- [15] Colin Ware. *Information Visualization: Perception for design*, Morgan Kaufman Publishers: San Francisco, CA, 2000.