Perceptual Considerations for Stereoscopy as a Variable Coding in Statistical Visualizations

Matthew Cutone
York University
Department of Psychology

5/6/2019

Introduction

Graphical displays and visualizations (i.e. a graph) utilize aspects of visual perception to encode abstract data into a form that can be incorporated into one's memory and cognition. Allowing the viewer to interpret and make inferences about the nature of the data.

Given that a large proportion of the human brain is involved in visual processing, the use of graphical visualizations utilizes the highest bandwidth channel for the passage of information from the environment to the observer (Ware, 2013). Properties of data are expressed as visual glyphs or symbols within the display, whose appearance and spatial relationship to other glyphs conveys meaning associated with each observation. For instance, in a 2-dimensional scatterplot presented on a computer screen or page, the Cartesian coordinate of a point on the graph representing a singular observation is determined by the magnitudes of its predictor and response variables. The spatial arrangement of points on the graph provides insight into the correlation between the variables, as well as differences between individual observations. However, a dataset may contain more than one predictor variable that can influence the outcome of an observation. If these additional variables interact with other predictor variables, this may result in peculiar patterns of points on a scatterplot such as clustering. Therefore, it is useful to include some representation for these additional variables in the display to explain such patterns.

On a 2D scatterplot, additional variables cannot be represented as displacements given the limited number of spatial dimensions. To overcome this constraint, the appearance of a glyph can be systematically varied along some feature dimension (e.g. colour, shape, size, orientation, etc.) to indicate additional categorical, qualitative, or quantitative variables associated with an observation. Bertin (2010) first described these visual features in the context information displays as "retinal variables" which provide perceptual encoding of data beyond what can be represented by the location of glyphs. The ability to use "retinal variables" takes advantage of the visual system's ability to bind such features together (Treisman & Gelade, 1980). Features are processed in stages requiring various degrees of attention, drawing upon greater cognitive resources and processing time. Glyphs in a plot which are similar along a single feature channel (e.g. color) are processed pre-attentively which causes target glyphs to 'pop-out' when searched for (A. Treisman, 1985; Ware, 2012). This allows the viewer to very rapidly identify cases within the dataset which are similar or distinct within the dimensions of some predictor variable. For instance, cases belonging to the same category labeled by colour are quickly localized within the field of the plot when searched for or perceptually grouped with similar glyphs. Furthermore, the glyphs of outlier cases can be coded in a way to be visually distinct (salient) from the bulk of the data, making them easy to identify pre-attentively. Due to the parallel nature of pre-attentive processing, the amount of time needed to search for items of interest among distractors does not increase much with numerosity of distractors (A. Treisman & Gormican, 1988). However, when two or more features occur in conjunction with another, such is the case when coding multiple variables as features into a single glyph (e.g colour and shape), more attentional resources must be drawn to bind features together. As a result, processing occurs serially which increases the amount of time needed to parse glyphs (A. M. Treisman & Gelade, 1980). This poses a challenge graphically presenting datasets with many predicator variables, where decoding and interpreting a plot becomes taxing to the viewer when multiple retinal variables are used to represent them. This can be exacerbated when viewing large datasets, given serial processing time multiplies with the number of elements present in the display. However, there are some cases where conjunctive-parallel processing of multiple features can occur at the pre-attentive level to emphasize targets (Driver, McLeod, & Dienes, 1992; DâZmura, Lennie, & Tiana, 1997; Theeuwes & Kooi, 1994; A. Treisman & Gormican, 1988; Ware, 2012). Furthermore, evidence suggests that stereoscopic depth in conjunction with color and motion facilitate pre-attentive processing (Nakayama, Silverman, et al., 1986). Therefore, evidence suggests that stereoscopic depth can be utilized to enhance the readability of glyphs encoding more than one predictor variable.

Visualizations using binocular disparity require specialized displays capable of presenting disparate images independently to each eye. Such displays are low-cost and utilize a variety of technologies suitable for different working environments. LCD shutter or polarized glasses, and mirror stereoscopes allow computerized displays to present binocular images by multiplexing or optically splitting incoming light between both eyes. More recently, consumer virtual and augmented reality headsets capable of stereoscopy are available. It is likely that these technologies may be used for data visualization tasks in the future as they are becoming increasingly prevalent in enterprise environments. Using stereoscopy and the added spatial dimension that comes with it increases the information density visualizations, however the efficacy of presenting data in such a way has yet to be examined comprehensively from an experimental perspective. The selection of studies presented in this paper provides psychophysical evidence showing potential benefits and issues surrounding the use of stereoscopy in contexts associated with statistical displays.

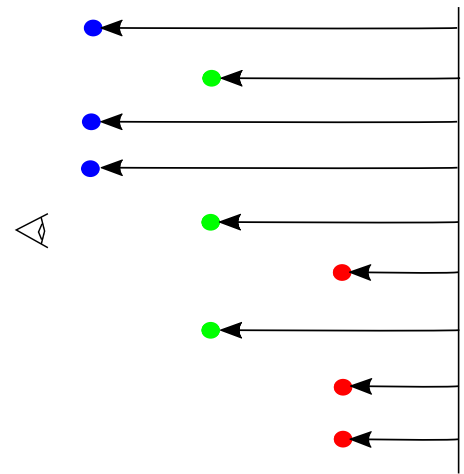

Figure 1. Example of using stereoscopic depth for coding a categorical variable. Points are displaced from the frontoparallel plane to a depth plane representing the coding variable. The eye represents the direction the viewer is looking towards the display. Glyphs here are redundantly coded with colour, however, colour can be varied to represent some other variable present in the dataset.

Why use stereoscopy in data visualizations?

The visual system uses retinal (binocular) disparity as a cue for depth (the percept is referred to as stereopsis, meaning “solid sight”), which is the interocular difference retinal images caused by their distinct viewpoints, where the magnitude of retinal offsets corresponds to an increase in perceived depth relative to the fixation point (Howard & Rogers). In addition to potentially improving the readability of plots as previously mentioned, depth from binocular disparity adds a spatial dimension for data to be presented. In a 3-dimensional scatterplot, glyphs are no longer constrained to horizontal (X) and vertical (Y) axes but can also be displaced in depth (Z) towards or away from the viewer. Furthermore, glyphs can be 3-D with variations of depth within the object. For instance, glyphs can be discs that can have a spatial location in three dimensions but may also be orientated along three axes of rotation, where the angle of each rotation can represent a different variable.

Humans evolved perceptual faculties to interact and interpret a world with three spatial dimensions, under normal circumstances, one can interpret the structure of the scene and safely interact with it. Therefore, one can assume that spatial perception efficacy should be equivalent across these dimensions. However, some research presented here shows that judgements of the spatial relationship between objects in depth are not necessarily veridical due to idiosyncrasies in depth perception, which may interfere with its effectiveness as a coding in statistical plots.

Pre-attentive processing and feature search

Nakayama et al. (1986)report that stereoscopic depth from retinal disparity does not require serial processing during search when presented in conjunction with colour and motion features. They propose that these features are decoded separately from disparity and their representations automatically duplicated across levels of disparities allowing for search to occur pre-attentively. Furthermore, searches occur quickly if objects share a common depth surface plane as people can switch attention between surfaces and search within them, especially if objects grouped by the surface vary along common feature dimensions (He & Nakayama, 1995). This suggests that stereoscopic depth can be leveraged as glyph feature encoding for variables in dense data visualizations. For example, a categorical predictor could be represented by varying the retinal disparity of a glyph, causing each category to fall on a different plane in depth (see Figure 1). The use of stereoscopic depth with glyphs in a graph will not interfere with parallel processing of other variables encoded by colour. One can also take advantage of the ability to attend to disparity defined surfaces and search within them. This may be useful where predictor variables are hierarchal, where each surface can index a level of a super-category, and glyphs resting on that surface are a subset of that category. This may provide a means for data to be subset perceptually without introducing another feature dimension to glyphs that may force the viewer to use serial conjunction searches. While this seems beneficial, it is important to note that shifting attention is known take longer in 3-D displays than 2-D ones (Atchley, Kramer, Andersen, & Theeuwes, 1997), so one must consider strategies to reduce costly comparisons between planes due to attentional switches.

Perceptual illusions

Stereoscopy can be used to attenuate perceptual illusions present in the visualization caused by the spatial organization of glyphs in the scene. The Muller-Lyer illusion is broken when the chevrons at the end of the line are displaced in depth (I. P. Howard, 2012), appearing to be spatially distinct from the line. In a variation of this illusion which may be encountered in data visualizations, the location of some target point in a cluster appears closer to another point in an adjacent cluster depending on where the points fall relative to the centroid of the cluster they belong to. This is due to the tendency of observers to judge the separation of the clusters based on their centroids, which in turn biases their judgment of the relative spatial positions of singular points between clusters (I. P. Howard, 2012). Harris and Morgan (1993) found that simply displacing a cluster in depth breaks the illusion, where their participants no longer made systematic errors attributed to the illusion. Furthermore, the Ebbinghaus illusion, where the perceived size of a shape is affected by the scale of shapes surrounding it is attenuated by displacing the target away from the plane of inducers (Papathomas, Feher, Julesz, & Zeevi, 1996). The use of stereoscopy provides a means of avoiding illusions which can affect the interpretation of the data. Given that stereoscopy seems useful for breaking illusions in clustered data, usually resulting from the presence of a categorical variable, it may be appropriate to redundantly code (Ware, 2012) categories with disparity and some other feature to improve visual comparisons of glyphs between clusters.

Stereoscopic contrast and averaging

Using stereoscopy as a coding for categorical or quantitative variables can create multiple impressions of form (I. P. Howard, 2012). Categorical coding where depths of glyphs are truncated to depth planes representing their factor level (Figure1) produce “depth transparency”, where one set of glyphs falls on a distinct plane relative to another, appearing as transparent surfaces overlaid on each other (Julesz & Johnson, 1968). For quantitative variables, where the predictor variable is expressed as the position of the glyph in depth along a continuous scale, can appear as a fuzzy (or snowy) volume of glyphs as they appear at various locations in depth (I. P. Howard, 2012). While these impressions are desired, if displays are dense (when visualizing large datasets for instance), psychophysical evidence shows that perceived depth of elements can change systematically.

When objects (points, lines, etc.) are spatially adjacent on the frontoparallel plane, yet have different depths due to binocular disparity, they are said to have depth contrast (I. P. Howard, 2012). Westheimer (1986) found that lines flanking a target appearing on different depth planes would attract or repel the perceived depth of the target line, depending on the angular separation of flanking lines. This suggests depth is “pooled” across neighboring elements when disparity contrast is locally high, whose affect is greater as the separation of flanking elements decreases (increased density). Westheimer (1986) found the critical separation for depth pooling to occur was 4-6 arcminutes. This phenomenon could be interpreted as a form of disparity averaging or normalization, where the disparity signal is smoothed to suppress noise in the disparity matching system (Fleet, Wagner, & Heeger, 1996; I. P. Howard, 2012). The effect of contrast was also observed when a horizontal target line was presented near adjacent points located above and below(Westheimer & Levi, 1987). This not only shows that pooling occurs in the vertical dimension, but even amongst objects with differing visual characteristics which may occur in plots.Mitchison and Westheimer (1984) found that closely spaced lines within a local reference frame altered the perceived depth of both the target and flanking lines, causing them to all to appear at different depths consistent with depth averaging. All lines appeared at the same depth when spacing was reduced further even though their disparity differed. This depth distortion and nullification effect due to adjacency of depth elements poses a problem for designers of visualizations, where the effect of a variable may be masked though contrast effects. This also poses a problem when providing reference marks in the display, for instance a grid of lines or points can be arranged to provide reference axes (so as not to have the glyphs floating in space with no context), or to define the origin of the plot in depth. The reference marks would have contrasting disparity with glyphs in the display, the resulting superposition with the glyphs may produce undesired changes in perceived depth. This effect was shown by Mitchison and Westheimer (1984) where a lattice of points in the background altered the perceived relative depth of the target lines, even in cases where the lines had no disparity.

Perception of stereo defined volumes

There has not been much research exploring the perception of stereo-defined volumes, as most studies have been focused on surfaces (Goutcher & Wilcox, 2016). However, volumes are of interest in the context of using stereoscopy for data visualization, as volumes would arise from the use of 3-D coordinates to encode variables. Glyphs coded in such a way would be spatially arranged corresponding to the statistical distribution they arise from. This means that centroid and density changes across the volume correspond to statistics such as mean, standard deviation, and range.

There have been attempts to use stereoscopy for data visualization in the past without first assessing its efficacy (Vogt & Wagner, 2011). Goutcher and Wilcox (2016) assessed the ability of participants to make judgments about the distribution (spread), location-in-depth, and range of stereoscopic volumes presented to them in relation to a reference. Volumes used were rectangular and defined only with points, whose distribution in depth was either uniform or Gaussian. Overall, they found that judgments were made using a limited amount of information presented in the stimuli. Participants seemed to rely on a simple model where judgments were produced using only the minimum and maximum disparity in the volumes presented. Furthermore, participants appeared insensitive to the shape of the distribution they were observing. Intuitively, the region in depth with highest density would correspond to the mean of the Gaussian sample. However, participants did not appear to use this fact when performing the task. Suggesting there is no automatic low-level faculty for doing so. Goutcher and Wilcox (2016) mention that with increased exposure duration, participants may be able to integrate additional information present in the display as seen in other research which could improve task performance. Note that their study used rectangular volumes whose distribution of points along the frontoparallel dimension (X-Y) were uniformly distributed. This is an unlikely configuration for data presented in a statistical display. Future studies may expand upon this research to examine how accurately people can estimate statistics from volumes that arise from statistical distributions across all spatial dimensions. Furthermore, one may examine whether people can estimate these statistics if multiple dimensions are correlated.

Conclusion

This paper reviewed several studies which are informative about the use of stereoscopy for generating data visualizations. Key points to consider are that stereoscopic depth as a feature dimension offers advantages in aiding pre-attentive visual search and information density on a display by adding an additional spatial dimension. On the other hand, filtering mechanisms related to disparity can affect the perceived appearance of glyphs in a display which can change their interpretation. Generally, there has not been substantial psychophysical research into the applied use of stereoscopy for presenting data and assessing how well viewers can extract statistical information from them. Whether these aforementioned factors play a significant role in the development and adoption of stereoscopy has yet to been seen.

Lastly, while taken for granted by most, a proportion of the population are “stereo-blind” where they cannot perceive depth from binocular disparity across their entire visual field (Richards, 1970). This is more complicated an issue to address than simply changing the colour swatch used in a plot as we do for colour-blind individuals. Even observers with good stereopsis may have difficulty perceiving depth accurately when presented on a display(McKee & Taylor, 2010) and may require a training regimen beforehand to ‘learn’ how to see depth (Fulvio & Rokers, 2017). Therefore, one must consider if the use of stereo displays is appropriate over more accessable and familiar formats.

References

Atchley, P., Kramer, A. F., Andersen, G. J., &

Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a âdepth-awareâ attentional focus. *Psychonomic Bulletin & Review*, *4*(4), 524–529.

Bertin, J. (2010). *Semiology of graphics: Diagrams, networks, maps*. ESRI PR.

Driver, J., McLeod, P., & Dienes, Z. (1992). Motion coherence and conjunction search: Implications for guided search theory. *Perception & Psychophysics*, *51*(1), 79–85.

DâZmura, M., Lennie, P., & Tiana, C. (1997). Color search and visual field segregation. *Perception & psychophysics*, *59*(3), 381–388.

Fleet, D. J., Wagner, H., & Heeger, D. J. (1996). Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision research*, *36*(12), 1839–1857.

Fulvio, J. M., & Rokers, B. (2017, nov). Use of cues in virtual reality depends on visual feedback. *Scientific Reports*, *7*(1). doi: 10.1038/s41598-017-16161-3

Goutcher, R., & Wilcox, L. M. (2016, sep). Representation and measurement of stereoscopic volumes. *Journal of Vision*, *16*(11), 16. doi: 10.1167/16.11.16

Harris, J., & Morgan, M. (1993). Stereo and motion disparities interfere with positional averaging. *Vision Research*, *33*(3), 309–312.

He, Z. J., & Nakayama, K. (1995, nov). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences*, *92*(24), 11155–11159. doi: 10.1073/pnas.92.24.11155

I. P. Howard, B. J. R. (2012). *Perceiving in depth, volume 2: Stereoscopic vision*. OXFORD UNIV PR.

Julesz, B., & Johnson, S. (1968). Stereograms portraying ambiguously perceivable surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, *61*(2), 437.

McKee, S. P., & Taylor, D. G. (2010, aug). The precision of binocular and monocular depth judgments in natural settings. *Journal of Vision*, *10*(10), 5–5. doi: 10.1167/10.10.5

Mitchison, G., & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research*, *24*(9), 1063–1073.

Nakayama, K., Silverman, G. H., et al. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*(6059), 264–265.

Papathomas, T. V., Feher, A., Julesz, B., & Zeevi, Y. (1996). Interactions of monocular and cyclopean components and the role of depth in the ebbinghaus illusion. *Perception*, *25*(7), 783–795.

Richards, W. (1970, aug). Stereopsis and stereoblindness. *Experimental Brain Research*, *10*(4), 380–388. doi: 10.1007/bf02324765

Theeuwes, J., & Kooi, F. L. (1994). Parallel search for a conjunction of contrast polarity and shape. *Vision Research*, *34*(22), 3013–3016.

Treisman, A. (1985). Preattentive processing in vision. *Computer vision, graphics, and image processing*, *31*(2), 156–177.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological review*, *95*(1), 15.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97–136.

Vogt, F., & Wagner, A. Y. (2011, sep). Stereo pairs in astrophysics. *Astrophysics and Space Science*, *337*(1), 79–92. doi: 10.1007/s10509-011-0801-z

Ware, C. (2012). *Information visualization: perception for design*. Elsevier.

Westheimer, G. (1986). Spatial interaction in the domain of disparity signals in human stereoscopic vision. *The Journal of physiology*, *370*(1), 619–629.

Westheimer, G., & Levi, D. M. (1987). Depth attraction and repulsion of disparate foveal stimuli. *Vision Research*, *27*(8), 1361–1368.