JENNIFER M. HOCKO

Information Visualization

The first section of this paper introduces two use cases for the attached Middle East prototype, and describes a few reasons why this information warrants a dynamic visual presentation. It also presents typical users' goals and describes some of the cognitive tasks users are currently expected to perform while searching for information about the Middle East. In the remaining sections, some common research themes associated with the creation of interactive, multidimensional visual representations are examined.

Use Case Descriptions and Rationale

As news reports deluge viewers with information about violence in the Middle East, this author frequently feels ignorant about 1) the countries in this region, as well as 2) the reasons behind current conflicts. With the assumption that many other people would also like to improve their knowledge and understanding of these topics, the Middle East prototype was constructed.

These topics warrant visualization for many reasons. First, there is a large quantity of data available about the Middle East (both in general and about particular conflicts), but interested users must currently search through multiple Web sites to find small portions of that information. Users are required to assemble disparate information about the Middle East in their minds (placing a high cognitive load on the user), rather than view the information in its entirety via an external medium (freeing users to absorb more information). Another difficulty is that in many cases, particular aspects about the Middle East (such as the availability of natural resources) may be closely linked to conflicts. This relationship and others like it cannot easily be construed from existing presentations. Finally, the current presentations (mostly Web pages full of text, with the occasional map here or there) do not take advantage of humans' advanced visual processing capabilities for perceiving correspondences, which if done, could allow for fast, deep, and long-lasting understanding of many complex data variables. As Card, et.al. suggest and support with their large collection of research articles, an effective information visualization using data about the Middle East could "amplify cognition" by "increasing resources, reducing search, enhancing recognition of patterns, [supporting] perceptual inference, and [providing a] manipulable medium" (pp. 16). For these reasons and many more, we should be able to use dynamic graphical aids to assist users with both their basic understanding of Middle East countries and in their understanding of this region's conflicts.

Dynamically Supporting Data Exploration Tasks

Several visualization experts and researchers advocate the use of dynamic, interactive displays for their ability to support data exploration tasks. Shneiderman even points to the usefulness of such displays for educational purposes like the one described in “Use Case Descriptions and Rationale” (pp. 239).

The benefits of dynamic displays for visualizing large, complex data sets in exploration-based situations are closely linked to their three primary characteristics: overview, zoom and filter, and details on demand. Card, et. al. combine the first and third characteristics into overview + detail, describing it as “two or more...linked visualizations” where one displays all the objects or provides a “visual framework,” and the other “shows a more detailed view of the [user-selected] object” (pp. 234). Zooming and filtering are the mechanisms by which users move between overview and detail; while the number of variables shown for each object increases, the total number of visible objects is consequently reduced (Card, et. al. pp. 234). Shneiderman indicates that “geographic applications are natural candidates for dynamic queries” with these characteristics, (pp. 237) and this author points to screen B of the Middle East prototype as an example. Users initially see an overview map of a country (in this case, Israel). However, when users move the box overlaid on that map, the Area of Detail changes to show a closer view of the selected region (such as mountains, valleys, rivers/oceans, and so on). Users can further tune that view using zooming and rotating controls. Lastly, the Data for Area of Detail provides general and/or specific data about the land, climate, and natural resources of the selected region by displaying content drawn from a database.

Ware notes that while employing these techniques does not provide simultaneous access to both overview and detail data, giving users the ability to easily and rapidly shift between both views supports their exploration goals more closely, and is therefore a key perceptual benefit of dynamic displays (pp. 360; also Fishkin and Stone, pp. 253; Keim, pp. 42). Further, the controls made available as part of dynamic displays guide users in constructing more effective queries, without exposing underlying technology requirements (Fishkin and Stone, pp. 253). These queries create the filter that determines which data is displayed and/or how it is displayed (Ahlberg and Shneiderman pp. 245). Because users can see—sometimes in a few tenths of a second—the effect their actions have on the visualization, they may also be “constrain[ed] from making erroneous or useless actions” (Ahlberg and Shneiderman pp. 245-246). If properly designed, the types of controls used may be so analogous to the user’s mental model of how things work in the physical world (for example, how a magnifying glass works to focus on a specific part of an object) that instructions may not even be necessary (Fiskin and Stone, pp. 257). Shneiderman also mentions that direct interaction with a data display gives users a more “playful” feeling when exploring a data set. Thus, dynamic displays may also increase users’ motivation for learning (pp. 239). Finally, since the query and system response are “tightly coupled,” the rapid feedback provided by the display may help users construct a mental model of how the information visualization works more quickly (Ahlberg and Shneiderman pp. 246). In sum, dynamic displays are effective in supporting data exploration tasks because they provide simple, easy-to-use controls that allow users to gather relevant information quickly, with less effort and likelihood of error.

Importance of Achieving the Pop Effect

In the context of building dynamic visualizations, various researchers explain the importance of achieving the “pop effect.” The pop effect is a term used to describe how certain visual elements of a display seem to automatically stand out while others do not. The pop effect has been explained by vision research, where it is scientifically labeled “pre-attentive processing.” Pre-attentive processing (also called “effortless” or “immediate” by Bertin) is described as the instantaneous decomposition of an image into its constituent objects, without the need to focus attention on those specific objects (Green, 3,4; 4.1). Ware describes “anything processed at a rate that is faster than 10 msec per item” as pre-attentive (pp. 165). Because the multiple objects that may comprise an image do not require individual attention, pre-attentive processing is also described as occurring in parallel. In contrast, attentive processing (called “effortful” or “sign-by-sign” by Bertin) requires viewers to consciously focus their attention on the individual objects that comprise an image. This scanning has also been described as “slow,” “arduous,” and “purposeful,” and is indicative of a serial type of processing (Green, 3.4; 4.1). The serial nature of attentive processing is supported by Ware’s assertion that “non-pre-attentive features are [processed at a rate of] 40 msec per item and more” (pp. 165).

Veluchkovsky et. al. explains that the pre-attentive processing stage of any visual search task “helps viewers locate objects in the visual world” while in contrast, the attentive processing stage “operates on a few objects at a time” and is thus “the bottleneck in visual processing” (pp. 79). Fekete and Plaisant further point out that attentive processing “will not scale well to displays [containing] millions of items [because] spotting a feature requires time linear to the number of features” (pp. 2). Ware believes that tasks like “area estimation judgements,” which might be a task for users of the Middle East prototype, can benefit from this rapid processing (pp. 168). It seems obvious that good visualizations would leverage pre-attentive processing to produce "information extraction at a single glance" (Green, 3.4; Ware, pp. 165), but the question is, how?

For pre-attentive processing to occur, a target image should differ from other images based on a single feature, such as color. Encoding an image using multiple features (such as color and orientation) requires viewers to encode and integrate the combination (or conjunction) of features, which in turn requires effort characteristic of attentive/serial processing (Green, 4.1; Ware, 168-169). Pre-attentive search seems to exist only "by examining the contents of a single feature representation," (Green, 4.1) while attentive search "involves searching for codings with a combination of features" (Ware, pp. 169). Ware describes many categories of pre-attentively processed features that are also discussed by the Gestalt theorists, including "form, color, motion, and spatial position" (pp. 165). Screen A of the Middle East prototype shows how color can be used to highlight selected information in a pre-attentive way. By using colors with a location on the visual light spectrum to which the eyes are particularly sensitive, designers can induce the pop effect. However, the most powerful feature for inducing pre-attentive processing is part of Ware's spatial position category (for reasons discussed in "The Mapping of Data Types to Visual Representations.") This power may also be understood in relation to the law of proximity, which states that "things that are close together are perceptually grouped together," and that viewers "perceptually group regions of similar element density" (Ware, pp. 203-204). For example, the relationships between objects in the Middle East prototype are enforced by their close proximity to one another. At the highest level, each screen (A,B,C, and D) contain objects that are related by a particular topic, such as "People" or "Conflicts." Proctor and Van Zandt indicate that "dials with common functions can be grouped together by varying factors such as proximity and similarity" (pp. 136). Therefore, at a more discrete level, controls that operate on particular objects are located close to the object they manipulate, such as the zoom and rotate controls for the map in screen B and the chart in screen C.

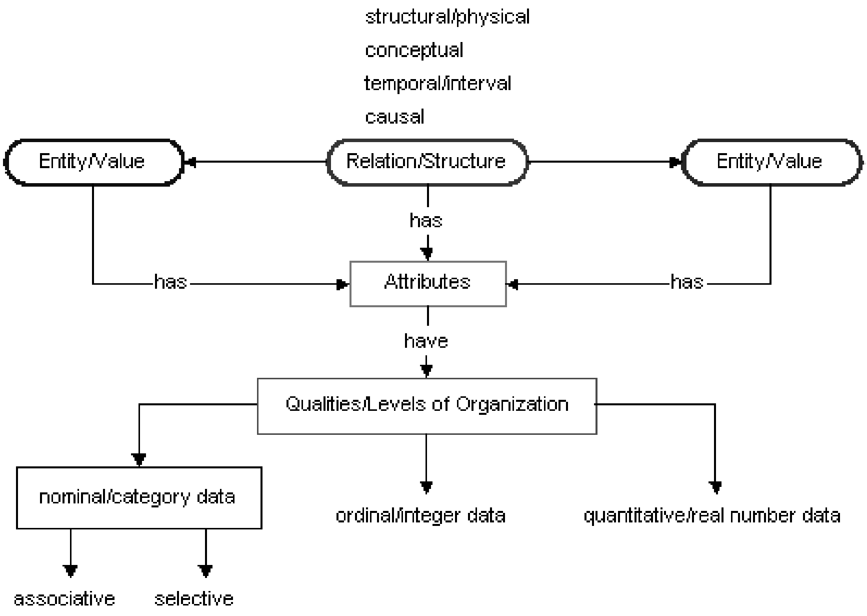
Being Selective About Encoding Techniques

Experts in the information visualization field agree that there is no single, generally-suited technique for optimally encoding all types of data. Rather, the way a variable is encoded (that is, what elements are used to produce an effective visual representation of it) depends on the variable itself (Keim, pp. 43). While a comprehensive taxonomy for individual data types and their ideal representations also does not exist, the several categorizations of variables (and the best ways for visualizing them) that have been presented can at least provide designers with a basis for creating effective displays (Ware, pp. 28). Most discussions on this topic focus on breaking variable information down into a number of discrete parts, doing the same with various visual elements (or "graphic primitives," as Bertin would call them), then attempting to create a map between the two. Ideally, such a map could be used not only to instruct designers, but also to create algorithmic languages that computer programs like Mackinlay's APT could understand and use to automatically render appropriate displays (pp. 66).

The Requirement to Categorize Knowledge into Data Types

Although Card, et. al., Ware (occasionally citing Stevens) and Green (citing Bertin) use different terminology, they basically break variable information down into the same categories. This breakdown is illustrated in Figure 1.

Figure 1 Variable Categorization



The entities or data values are the "objects of interest," while the relations or data structures illustrate the relationship between entities (Ware, pp. 28, 29). There are many different types of relationships, including structural/physical, conceptual, temporal/interval, and causal. Further, “both entities and relationships can have attributes,” which also have different qualities (Ware, pp. 29). Not surprisingly, Ware’s descriptions of attribute quality coincide nicely with Bertin’s “levels of organization,” or the “type of data scale” for a given variable (pp. 29; Green, 3.3.2). Nominal levels (or “category data”) are essentially unordered sets of information, which can be further classified into associative (where group membership is clear) and selective (where the variable can be easily identified as separate from surrounding objects (Green, 3.3.2; Ware, pp. 30). Ordinal levels are ordered sets or integer data, and quantitative (real-number data) are characterized by numeric ranges or straight-forward quantities that “permit direct extraction of ratios” (Green, 3.3.2; Ware, pp. 29-30).

The Examination of Visual Variables

Bertin’s *Image Theory* was the first to assert that every visualization is comprised of some number of “graphic primitives,” which are grouped into “invariants, components, correspondences, and marks.” The relations [correspondences] previously described are “depicted by mapping the components [that is, the entities] to some visual graphic variable and placing a ‘mark’ on the visualization” (Card, et. al., pp. 28; Green, 3.1). Bertin further classifies these “visual variables” into planar (spatial) and retinal (object); the retinal variables being “size, color, shape, orientation, texture, and brightness” (Card, et. al., pp. 29-30; Green, 3.2) and named for their unique ability to stimulate the retina “independent of [their] position” (Card, et.al., pp. 29). Interestingly enough, the retinal variables Bertin cites are very similar to Ware’s “features that can be pre-attentively processed (pp. 165)” and Proctor and Van Zandt’s “coding dimensions” (pp. 170), and have much in common with the Gestalt theories (Ware, pp. 203-213).

The Mapping of Data Types to Visual Representations

Bertin attributes the design of both ineffective and error-prone visualizations to a mismatch between the information to be visualized and the visual elements selected to encode that information. In other words, the mapping between variables and graphic primitives is often incorrect (Green, 3.3). Thus, Bertin (and others like MacEachren) provide advice about how to choose an appropriate planar or retinal variable by describing which visual variables best match which data types (Card, et. al., pp. 30). (See the attached Data Table for an example mapping that was used to create the Middle East prototype.)

While a complete discussion cannot be included here, the most important aspect of Bertin's advice—which has since been proven through empirical study—is the unique capability of the planar (spatial) visual variable to effectively encode every data type described in "The Requirement to Categorize Knowledge into Data Types." Card, et. al. indicate that "space is perceptually dominant," and "is such a good visual coding of data that the first decision of visualization design is which variables get spatial encoding at the expense of others" (pp. 26). Bertin apparently went so far as to say that "error cannot occur when using the planar variables" (Green, 3.3.2). There are several theories about why space holds this key perceptual position: some describe the brain as "a collection of quasi-independent modules," each of which contain neurons tuned to spatial image features (while the retinal variables, such as color, only stimulate neurons in one module) (Green, 4.0); others like Attneave believe that "spatial location is the attribute which holds retinal features together when time comes for reassembly [of an image]" (Green, 4.0); others still, like Zeki, theorize that there are two systems in the visual cortex specifically dedicated to perceiving form (Wright, pp. 85).

The Psychophysical Scaling Explanation

In addition to the different categories of data that they can effectively express, visual variables also have different capabilities for conveying length. Green defines length as "the number of [clearly distinguishable] categories or steps" that a variable can convey, giving an example of brightness levels (3.3.1). If the length of a variable does not directly map to the scale of data values requiring visualization, perceptual problems are likely to occur (Green, 3.3.1).

While Bertin did provide some rules for determining the length of a given variable, empirical research into psychophysical scaling provides more scientific explanations. For procedures that use "direct scaling," observers are asked to measure the intensity of a specific stimuli, such as the brightness of a color or the loudness of a sound. Based on a value that is assigned to one stimulus as a standard, the "magnitude of other stimuli [is rated] in proportion to the standard" (Proctor and Van Zandt, pp. 70). These studies show that "a doubling of intensity does not produce proportional [that is, linear] increases in sensation. [Rather,] a doubling of intensity is perceived as a far smaller increase in brightness [or other stimulus]" (Green, 4.2.1). Therefore, determining a pre-attentive "just notable difference" advocated by information designers like Tufte may in fact be limited by the inherent qualities of the visual variables that are available (pp. 73; also Green, 4.2.2)

Investigating Techniques for Overcoming Perceptual Barriers

The goal of all effective and expressive visualizations is to present a large amount of complex data in a manner that leverages the strengths of our visual perceptual systems and to some degree, reduces its limitations. One main limitation that information visualizations must overcome is that of visualizing data with more than 3 quantitative variables. Bertin's original mapping of data types to visual variables illustrates only two ways to visually represent quantitative data: by using the planar (spatial) variables (which provide both x and y coordinates) and the retinal variable of size (Green, 3.3.2, Table II). If one were to go beyond this mapping and introduce another retinal variable, for example, the onus would be on the viewer to integrate those features, resulting in the slow, attentive, conjunction search described in "Importance of Achieving the Pop Effect." Therefore, Bertin believed that "efficient" (that is, pre-attentively processed) visualizations are limited to three components, and introduced what still appears to be a barrier for many visual representations (Green, 3.4).

Fortunately, Bertin failed to take into consideration two visual variables by which conjunctive searches seem to be pre-attentively performed: biocular disparity and motion (Green, 5.4.1). Biocular disparity is a term used to describe how the two, slightly different views received by each eye are merged together into a single image, and is relevant to discussions about depth perception (Proctor and Van Zandt, pp. 140). The ability of "disparity [to provide] accurate information about depth relative to [a] fixated object" (Proctor and Van Zandt, pp. 141) implies that if properly executed, humans could efficiently process 3-dimensional displays; in other words, three planar variables (x, y, and z axes) rather than just Bertin's two (x and y axes). However, because the use of 3-D visualizations present new problems such as distortions and occlusion (Ware, pp. 287), researchers like Wright strongly urge that such displays not be used without the additional support provided by "motion and animated interaction" (pp. 83). Ware (and others like Green) are now recognizing that "we can treat motion as an attribute of a visual object much as we consider size, color, and position as object attributes" (pp. 232) and that "humans are reasonably sensitive to motion" (pp. 234). Thus, a combination of biocular disparity and motion (such as that suggested in screens A and B of the Middle East prototype) may help careful present-day designers overcome some limitations of displaying more than 3 quantitative variables.

References

Card, S.K., Mackinlay, J.D., and Shneiderman, B., ed. (1999). Readings in Information Visualization: Using Vision to Think. San Francisco, CA: Morgan Kaufmann Publishers, Inc.

> Ahlberg, C. and Shneiderman, B. “Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays,” pp. 244-250.
>
> Fishkin, K., and Stone, M.C. “Enhanced Dynamic Queries via Movable Filters,” pp. 253-259.
>
> Mackinlay, J. “Automating the Design of Graphical Presentations of Relational Information,” pp. 66-82.
>
> Shneiderman, B. “Dynamic Queries for Visual Information Seeking,” pp. 236-243.
>
> Wright, W. “Research Report: Information Animation Applications in the Capital Markets,” pp. 83-89.

Fekete, J-D. and Plaisant, C. (2002). Interactive Information Visualization to the Million. University of Maryland, College Park, MD: Human Computer Interaction Lab.

Green, M. (1998). Toward a Perceptual Science of Multidimensional Data Visualization: Bertin and Beyond. http://www.ergogero.com/dataviz/dviz0.html. Accessed 9 April 2002. ERGO/GERO.

Keim, D.A. (2001). Visual Exploration of Large Data Sets. Communications of the ACM,(44)8, pp. 39-44. New York, NY: ACM Press.

Proctor, R.W., Van Zandt, T. (1994). Human Factors in Simple and Complex Systems. Needham Heights, MA: Allyn and Bacon.

Tufte, E.R. (1997). Visual Explanations: Images and Quantities, Evidence and Narrative. Cheshire, CT: Graphics Press.

Velichkovsky, B.M., Dornhoefer, S.M., Pannasch, S., and Unema, P.J.A. (2000). In Proceedings of Eye Tracking Research & Applications Symposium 2000, pp. 79-85. New York, NY: ACM Press.

Ware, C. (2000). Information Visualization: Perception for Design. San Francisco, CA: Morgan Kaufmann Publishers.