Visualization Techniques from Design & Cartography

Robert Simmon

Abstract

Data visualization has a long history, dating back to the origin of writing, while the history of cartography spans several thousand years. During that time both fields developed effective approaches to display information clearly. Unfortunately, modern computer-based visualization practices often neglect this history. Many data visualizations can be improved by incorporating traditional techniques of visualization and modern design methods based on the knowledge of human perception.

Introduction

The history of data visualization is intertwined with the history of communication. Cuneiform, the earliest form of writing, evolved out of a system of pictograms used to keep track of stores and taxes in Sumerian city-states (Meggs, 1998). The pictograms were representative symbols used to indicate the types of stored goods, combined with a decimal numbering system that denoted quantity. Over time the pictograms became less literal, and eventually transformed into abstract cuneiform script. The need to record and communicate numerical data triggered the development of written language.

The earliest maps were also developed in the Middle East, with early surviving examples drafted by ancient Babylonians and Egyptians (Wikipedia ‘History of Cartography,’ 2009). By the time of the Roman Empire maps appeared that showed networks and connections instead of simply representing land surfaces. One example is the Tabula Peutingeriana, a medieval copy of a 4th century map of the Roman public roads. The roads linking settlements and towns in the ancient world are drawn on a 6.75-meter-long scroll (Wikipedia ‘Tabula Peutingeriana,’ 2009). Types of locations are differentiated by standardized symbols. Distances between locations along the road network are emphasized, rather than their exact geographic relationships. These early map-makes knew one of the key principles of information design: carefully highlight vital data.

Further developments in thematic maps and data visualization occurred during the Enlightenment. Edmund Halley published a map of the trade winds in 1686. Wind speed and direction were indicated with tapered lines. This was one of the first maps to show quantitative data rather than geographical relationships (Denis and Friendly, 2001–2008). William Smith’s 1815 map: *A Delineation of the Strata of England and Wales, with Part of Scotland*, was the first large-scale geological map (Winchester, 2001). Smith used color to denote different rock types that lay under the British countryside, a qualitative thematic map. *A Delineation of the Strata of England and Wales* was both large (approximately 6 by 9 feet) and detailed, and simultaneously provided an overview of the geology of Britain and details of specific locations. Neither Halley or Smith worked alone. Scientists of the times collaborated with teams of printers, engravers, and other craftsmen who adapted the data to the media. The results were elegant, clear, and effective.

In the 1950s and 60s computer graphics were born, which would begin a revolution in the display of information. Early displays were very small, monochrome, and low resolution.

Full-color displays did not become common until the late 1990s, and screen resolutions (roughly 100 dots per inch) are currently much lower than print (300-2400 dots per inch). For data visualization, computers had a major advantage over tradition production methods: scientists could visualize their own data. This enabled scientists to produce graphics quickly and inexpensively, but designers were now rarely involved in data visualization. Many of the presentation decisions are made by programmers writing commercial software applications, which are often designed to optimize performance or ease programming rather than to produce effective graphics. As a result many valuable graphics techniques—developed over centuries—are neglected in scientific visualization.

A Definition of Information Design

> …addresses the organization and presentation of data: its transformation into valuable, meaningful information.
>
> Nathan Shedroff
>
> The art and science of preparing information so that it can be used by human beings with efficiency and effectiveness.
>
> Robert E. Horn

(Visocky-O'grady, 2008)

Information design is a multi-disciplinary field that combines elements of cognitive science, graphic design, and statistical graphics. Practitioners aim to display information clearly with an emphasis on the communication of ideas. Some simple information design principles can transform hard-to-read, unappealing illustrations into figures that are clear and attractive.

Case Study: Map of Sea Ice Age

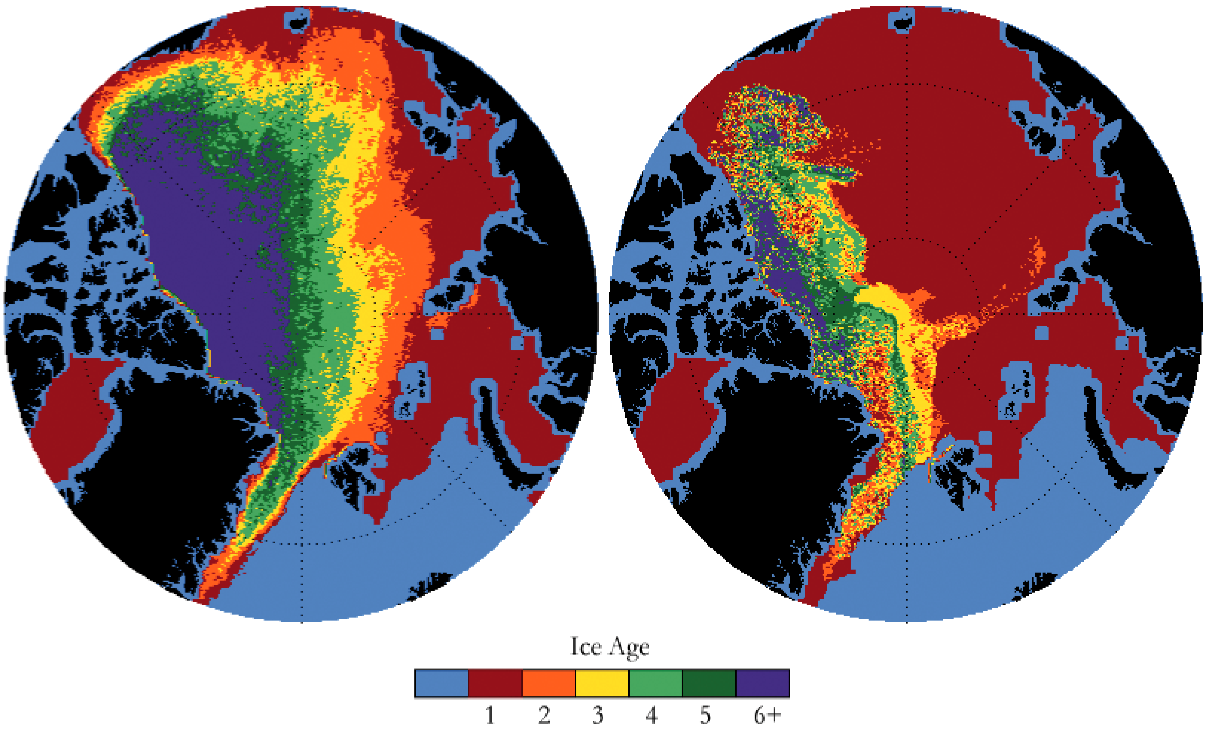

Figure 1. These maps compare the average age of sea ice in the Arctic ocean from 1985-2000 to the age of the ice in February 2008. Although it is possible to distinguish multi-year ice from fresh (one-year-old) ice, the design of the maps makes it more difficult than necessary to interpret the patterns of the ice. (Image courtesy National Snow and Ice Data Center.)

This pair of maps shows the age of sea ice in the Arctic (Cook-Anderson, 2008). On the left is the mean age of ice in February during the years 1985 to 2000; on the right is displayed the age of sea ice in February of 2008. Several of the design choices of the mapmakers obscure patterns in the data, rather than emphasizing important elements.

Graphics for static, low-resolution screen display should be *anti-aliased.* Because the resolution of computer displays is much lower than the ability of the human eye to perceive detail, curved edges and diagonal lines appear stair stepped (below, left) if solid colors are used. Lines and edges with blended colors, however, appear smooth (below, right).

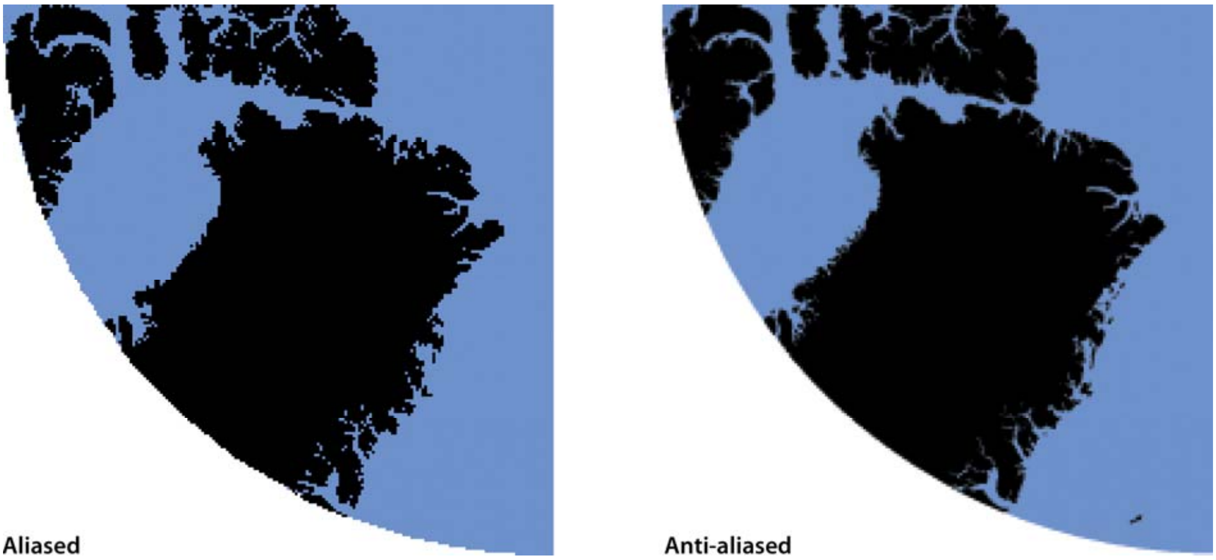

Figure 2. Anti-aliasing is a technique used to smooth curved edges in low-resolution displays. In the original aliased image (left) pixels are either land or ocean, leaving a coarse appearance to the coastline. Applying anti-aliasing to the land mask in this map (right) shows more detail than the original, and reduces the amount of visual noise in areas with high-frequency detail. (Image courtesy Robert Simmon, Sigma Space/NASA Goddard Space Flight Center.)

Another important element of information design is to emphasize the data, rather than ancillary or irrelevant information (Tufte, 2001). One method of emphasizing data is to place it in the foreground of an image. Design elements that are bold, heavy, saturated, and sharp stand out as foreground elements. In the original sea ice age map the ice-free ocean and land areas are visually balanced with the ice-covered areas, so a viewer does not automatically focus on the data. A better approach is to use light, de-saturated colors for areas of no data, while using saturated colors for data. This will provide contrast between data-filled portions of an image and less important information.

Many maps use color to represent numerical values at points in space. Colors and gradients should be chosen to be compatible with human vision so that changes in data are proportional to perceived changes in quantity. Unfortunately one common color scheme, the rainbow (or spectral) palette, distorts the underlying data (Rogowitz and Treinish, 1996). (The rainbow palette is a gradient that progresses through the hues of the rainbow: red, orange, yellow, green, blue, indigo, and violet.)

Figure 3. The "rainbow" palette is a common color scheme used for coding numeric values as colors in data visualization. Unfortunately this palette is difficult or impossible for the human visual system to interpret properly. Some color transitions (blue to green and yellow to red, for example) occur very quickly, which introduces false contrast, while other colors (green in particular) cover wide ranges of the palette, obscuring detail. (Image courtesy Robert Simmon, Sigma Space/NASA Goddard Space Flight Center.)

The rainbow palette has several flaws. First, perceived changes in value are nonlinear. Some regions of the scale appear to change rapidly (the transitions from blue to cyan, green to yellow, and yellow to red), while other regions appear to change very slowly (green). This effect occurs because the red, green, and blue primaries displayed by computer screens are not perceived as equally bright by humans. Secondly, the brightness of the palette is not ordered. (Ware, 2004) Color schemes that vary in either luminance (dark to light) or saturation (dull to bright) have a natural sequence, while those that vary in hue (red to violet) do not.

Perceptual color spaces like Munsell or CIE L*a*b can be used to generate color palettes that display data accurately (Brewer, 2002). Palettes should be designed in the perceptual space, and then mapped to the RGB values of computer displays. One compromise that preserves relationships between data points and allows viewers to accurately read data values is to use a palette that varies uniformly in brightness, while simultaneously changing both hue and saturation. The change in brightness conveys sequential information most strongly, while the shift in hue and saturation helps a viewer distinguish colors from one another. Color-blind viewers (8% of males) can also accurately read these maps, and they retain much of their accuracy when printed in black and white or photocopied.

A special type of color palette, described as a *divergent* palette (Brewer, 2002), is helpful in the display of quantities with positive and negative values, or that vary from a mean. These palettes are centered on a neutral color, and ramp to two equally luminous and saturated values to represent extreme values of equal magnitude. In these palettes, positive and negative values are almost instantly identifiable, and quantities of equal absolute value are equally prominent.

This redesigned map of multi-year Arctic sea ice uses these principles. I have anti-aliased the boundary between land and water, lightened the land area and ice-free water considerably, while applying a perceptually based color palette to the sea ice age data. The oldest ice is dark and saturated, which makes it stand out from the background areas, and patterns in the ice are more clear. There is a slight sacrifice in the ability to interpret the age numerically, but the purpose of the map is to show the drastic decrease between the historical winter ice coverage in the Arctic and the conditions in February 2008.

Data visualization allows viewers to survey and interpret quantitative information quickly and easily. Perception-based principles from graphic and information design enhance these qualities of visualization, and improve the ability of graphics to convey information.

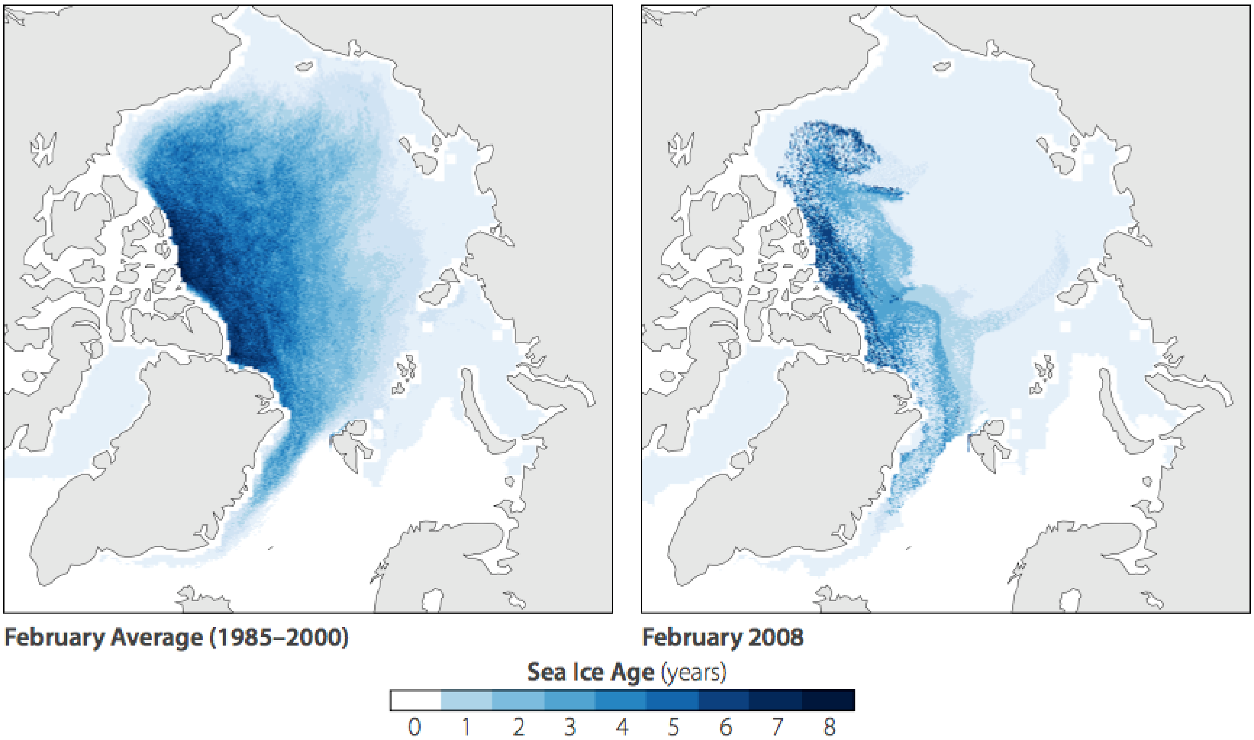

Figure 4. These sea ice maps use simple design principles to enable rapid interpretation of the data and show patterns clearly. The techniques include anti-aliasing edges, reinforcing the contrast between the data and the background, and displaying ice age using a perceptual color palette. (Image courtesy Robert Simmon, Sigma Space/NASA Goddard Space Flight Center, based on data from the National Snow and Ice Data Center.)

References

Brewer, Cynthia (2002), 'Color Brewer', http://www.personal.psu.edu/cab38/ColorBrewer/ColorBrewer_intro.html, accessed 14 September 2009.

Cook-Anderson, Gretchen (2008), 'Researchers Say Arctic Sea Ice Still at Risk Despite Cold Winter', http://www.nasa.gov/topics/earth/features/seaice_conditions_feature.html, accessed 15 September 2009.

Denis, Daniel J., and Friendly, Michael (2001–2008), 'Milestones in the History of Thematic Cartography, Statistical Graphics, and Data Visualization', http://www.math.yorku.ca/SCS/Gallery/milestone/index.html, accessed 14 September 2009.

Meggs, Philip (1998), *A History of Graphic Design, 3rd Edition*, John Wiley & Sons.

Rogowitz, Bernice and Treinish, Lloyd (1996) 'Why Should Engineers and Scientists Be Worried About Color?' http://www.research.ibm.com/people/l/lloydt/color/color.htm, accessed 14 September 2009.

Tufte, Edward (2001), *The Visual Display of Quantitative Information, Second Edition*, Cheshire, CT: Graphics Press.

Visocky-O'grady, Jenn (2008), *The Information Design Handbook*, How.

Ware, Colin (2004), *Information Visualization, Second Edition: Perception for Design*, Morgan Kaufmann.

Wikipedia (2009), 'History of Cartography', http://en.wikipedia.org/wiki/History_of_cartography, accessed 14 September 2009.

Wikipedia (2009) ‘Tabula Peutingeriana,’ http://en.wikipedia.org/wiki/Tabula_Peutingeriana, accessed 14 September 2009.

Winchester, Simon (2001), *The Map That Changed the World: William Smith and the Birth of Modern Geology,* HarperCollins.

Biography

Robert Simmon is a data visualizer and designer for NASA’s Earth Observatory web site. With 14 years of experience at NASA, he is an expert at creating clear and compelling imagery from satellite data. He focuses on producing visualizations that are elegant and easily understandable, while accurately presenting the underlying data. His imagery has appeared in newspapers, web sites, and advertisements, and is featured on the login screen of the Apple iPhone.