

A Data Visualization Case Study: Ash from Eyjafjallajökull Volcano

Robert Simmon

Sigma Space Incorporated, under contract to NASA Goddard Space Flight Center. Code 613.2. Greenbelt, MD 20771, USA
— robert.simmon@nasa.gov

Although data visualization is a powerful tool in Earth science, the resulting imagery is often complex and difficult to interpret for non-experts. Students, journalists, web site visitors, or museum attendees often have difficulty understanding false-color imagery and data-driven maps. Many visualizations are designed for data exploration or peer communication. Different techniques are necessary for communication with a broad audience. Data visualization combines ideas from cognitive science, graphic design, and cartography, and applies them to the challenge of presenting data clearly. Visualizers on NASA's Earth Observatory web site team use these techniques to craft images based on remote sensing data. Imagery of the eruption of Iceland's Eyjafjallajökull Volcano serves as a case study, showing specific applications of general design techniques. By using color carefully, precisely aligning disparate data sets, and highlighting important features, we crafted an image that clearly conveys the complex distribution of airborne ash.

Keywords: Visualization, design, color, communication, pre-attentive processing.

1. Introduction

Images are one of the most effective ways to communicate data-rich scientific information (Tufte, 2001). Well-crafted visualizations can accurately show large datasets and convey complex concepts, even to non-specialists. Unfortunately, most scientific visualizations are created for data analysis and peer communication, and are not ideal for public outreach. This occurs for several reasons: few scientists receive training in visualization, data graphics often follow discipline conventions rather than cognitive principles, and the software typically used to create imagery can make it difficult to precisely control graphical output. Ideas from four disciplines—cartography, information visualization, graphic design, and cognitive science—would aid scientists in developing more accurate, appealing, and informative imagery.

I used some of these ideas in a data visualization of ash from Iceland's Eyjafjallajökull Volcano over northern Europe, published on NASA's Earth Observatory web site (NASA, 2010). The airborne ash image specifically illustrates several concepts of data visualization: perceptually based color maps, the importance of context, and pre-attentive processing.

2. Background

For several days in April of 2010, air traffic in Europe was almost completely shut down by ash from Iceland's Eyjafjallajökull volcano. The widespread flight cancellations weren't caused by ash filling the skies all over Europe, but by uncertainty in the ash's location. Without knowing exactly where the ash was, air traffic controllers couldn't risk allowing passenger flights to embark.

Currently, ash forecasts are based on computer models that predict the movement of volcanic ash based on its observed location and altitude, combined with wind speed and direction. In the case of Eyjafjallajökull, the initial location of the ash was known (the volcano's summit), but not the altitude. This initial uncertainty grew as the ash blew towards Europe, dispersing and moving up and down in the atmosphere. The only way to constrain the computer models' forecasts was to observe the ash. Comprehensive observations are impossible to make from ground-based instruments, which are widely scattered, and it is difficult and dangerous to directly sample in the air. However, flying in space, far above the ash, some satellites can track ash even as it spreads over long distances. Forecasters can use measurements from these satellites to improve predictions of ash movement, reducing the amount of airspace closed to flights.

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite uses reflected light from a laser beam to measure particles in the atmosphere. Critically for ash forecasting, it measures both the location *and* altitude of particles. It can also see ash within clouds. CALIPSO observations are limited, because it only measures aerosols in a line directly underneath the satellite.

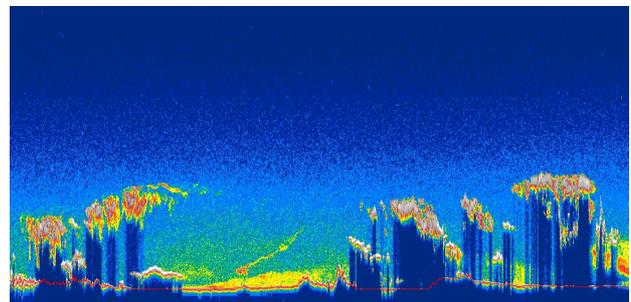


Figure 1: CALIPSO aerosol profile, May 16, 2010.

On the Earth Observatory, we wanted to show CALIPSO data (Figure 1) to complement the large number of visible light, photo-like images we'd acquired, and to advertise the data, which is considered experimental. The nature of the data presented an interesting visualization challenge: it's a two-dimensional curtain stretching from the Earth's surface to the stratosphere, along the satellite ground track. Simply displaying an automatically generated browse image is not very informative. To improve the graphic, we modified the color palette, and added imagery showing the horizontal distribution of ash and underlying geography.

3. Color

Color selection in data visualization is not merely an aesthetic choice; it is a crucial tool to convey quantitative information (Tufte, 1997). Properly selected colors convey the underlying data accurately, in contrast to many color schemes commonly

used in visualization that distort relationships between data values.



Figure 2: Original CALIPSO color palette.

Figure 2 shows the original palette used for the display of CALIPSO aerosol profile data. The differences between colors vary sporadically, with some transitions abrupt (blue to cyan, pink to gray), and others gradual (yellow, orange, red). Furthermore, lightness in the palette does not change monotonically. Even with steadily increasing data values the colors vary from dark to light to dark to light in an uneven progression. This is demonstrated in figure 3, which is the color CALIPSO palette converted to a grayscale image. Because of this unevenness, a viewer cannot easily discern patterns in the data.



Figure 3: CALIPSO color palette converted to grayscale.

3.1 Color Perception

Humans perceive color in three dimensions. Not the red, green, and blue used to define colors on computers; but lightness, hue, and saturation (NASA Ames Research Center, n.d.). *Lightness* is simply a continuum from black to white. *Hue* is determined by the wavelength of light, and is colloquially referred to as color. It varies from red to orange, yellow, green, blue, indigo, and finally violet. *Saturation* is the intensity of a color. Intense colors are highly saturated, while earth tones are desaturated. Black, gray, and white are unsaturated. Figure 2 shows these axes. Each bar varies one parameter at a time, leaving the other two constant.

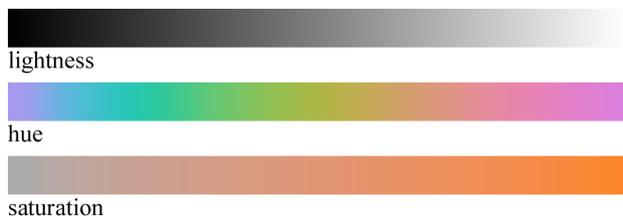


Figure 4. The three axes of color perception.

Of these three dimensions, lightness is perceived most strongly, followed by hue and saturation. Because of this, color palettes in data visualization should be based on lightness, with hue and saturation used as secondary components. A palette that varies linearly in lightness will be perceived more uniformly than one that is nonlinear, or changes sporadically from light to dark. Variations in hue and saturation reinforce the change in lightness, and help mitigate simultaneous contrast (the phenomenon where the perceived color in a given area of an image is modified by nearby colors). Palettes that vary monotonically in lightness also reproduce accurately in black and white, and can be accurately read by colorblind viewers.

Carefully chosen colors that correspond to cultural norms and everyday experiences can aid image interpretation for non-experts (Phipps, 2010). For example, a beige-green palette for vegetation, or blue for cold and red for hot. On the other hand,

arbitrary color scales, such as the rainbow palette, often confuse novices.

3.2 Perceptual Palettes

To accurately show spatial patterns in data it is important that the changes in lightness, hue, and saturation be perceptually linear (Rogowitz, 1993). For any change in the value of the data being represented, the *perceived* change in the palette should be equivalent: a change from 5 to 10 should appear to be the same as a change from 45 to 50. (Logarithmic scales should be proportional: the change from 0.1 to 1 is equivalent to the change from 10 to 100.)

Unfortunately, the most common tools for creating palettes are not optimized for creating perceptual palettes. They are based on the representation of color by computers: 3 channels [red, green, and blue (RGB)], each determined by 8 bits of data. Since the wavelengths of light emitted by the red, green, and blue pixels on a computer screen do not exactly correspond to the sensitivity of the retina's color-sensitive cells, the three channels aren't seen to be equally bright. Therefore, equal changes in red, green, and blue color space are not perceived as equal changes by an observer. Even color pickers based on hue, saturation, and brightness (HSB) are determined by a simple remapping of red, green, and blue channels.

Perceptually-based palettes must be created in a color space that accurately represents human vision, such as CIELAB or Munsell. Two web resources for creating these palettes are Color Brewer (<http://colorbrewer2.org/>) and the NASA Ames Color Tool (<http://colorusage.arc.nasa.gov/ColorTool.php>).

The color scale we used to show Eyjafjallajökull ash detected by CALIPSO was derived from the Color Brewer "Oranges" palette (Brewer, n.d.). It varies from burnt umber, to orange, to light peach (Figure 5). Lightness and saturation change linearly while hue is constant. The orange tints are meant to intuitively represent the concept of aerosols.



Figure 5: Eyjafjallajökull ash color palette.

4. Context

Audiences unfamiliar with a particular data set often need context. In isolation, the aerosol profile data did not show the relationship between the ash and the areas of Europe affected by flight cancellations. We needed to show the horizontal distribution of ash in addition to the vertical distribution. Unfortunately, the data were acquired at night—so we couldn't simply use a natural-color image. Instead, we needed to rely on thermal infrared data, which is essentially a measure of temperature. Volcanic ash and clouds (at the same altitude) will almost always have the same temperature, and will look the same in thermal infrared imagery. In addition, a thin ash plume may be completely invisible in thermal infrared wavelengths.

Ash, however, emits thermal infrared radiation *slightly* differently than water, so two images taken at different wavelengths will appear different from each other where there is ash, but not where there are clouds (Ellrod, 2010). This is called the *split window* technique. Observations coincident with the CALIPSO overpass were available from the Moderate Resolution Imaging Spectroradiometer (MODIS), which

collects data in several thermal infrared bands appropriate for calculating a split window. Subtracting the 11 μm band from the 12 μm band yields the grayscale image shown in Figure 6.

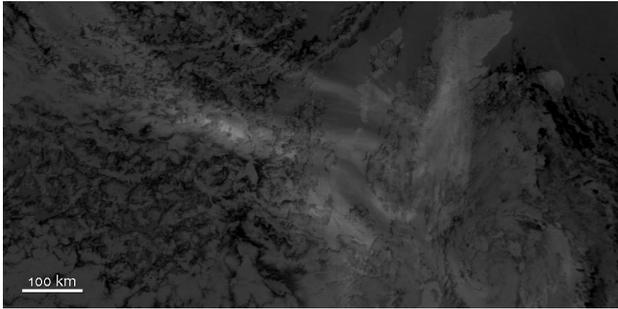


Figure 6: Grayscale split-window image of ash distribution.

By itself, the split-window image is not very informative: for the most part ash merely appears slightly lighter than the background, whether it is ocean, cloud, or land. We needed to develop a technique to better distinguish ash from the background. Our first attempt was to combine the split window image (12 μm minus 11 μm) in the red channel with the original bands, 11 μm and 12 μm in the green and blue channels. The result was more colorful, but not much more informative (Figure 7).



Figure 7: Multi-band composite image of split-window and thermal infrared data.

Other band combinations also failed to show the ash clearly. They are understandable to experts, but not to novices. An easily understandable image needed to be somewhat familiar (i.e. white clouds), and the ash needed to clearly stand out from the background. Our solution, shown in Figure 8, was to combine a colorized split window image (which emphasized the ash) with an inverted copy of one of the thermal infrared channels (which revealed the clouds). The ash appeared as distinct yellow areas on a background similar to the satellite images shown on television weather forecasts.

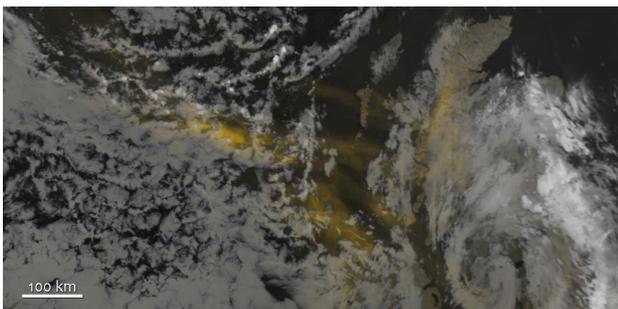


Figure 8: Final image of ash extent.

5. Preattentive processing

The ash stands out from the background because the visual system is highlighting the yellow areas of the image before the conscious brain becomes aware of the image—the visual system is focusing conscious attention to the yellow areas. This is known as preattentive processing (Ware, 2004). Preattentive processing refers to features that stand out from the background, that are immediately visible. In typography, words that are **bold** or *italic* are processed preattentively, which is why they are useful for emphasis. According to Ware, the following features are processed preattentively:

Form: line orientation, line length, line width, line collinearity, size, curvature, spatial grouping, blur, added marks, and numerosity.

Color: hue, intensity.

Motion: flicker, direction of motion.

Spatial Position: 2d position, stereoscopic depth, convex/concave, shape from shading.

Through manipulation of these properties, a designer can guide the attention of a viewer. Preattentive processing can be used to create a visual hierarchy, highlight a particular data point or feature, and group similar items together. Knowledge of preattentively processed features also allows the creation of multivariate graphics, with each variable encoded by a separate feature.

6. Conclusion

Accurate, attractive, and clear imagery is a key component of science communication to non-experts. Scientists can create better visualizations by adopting techniques from cartography, information visualization, graphic design, and cognitive science. A case study using coincident, multi-instrument observations of ash from the Eyjafjallajökull Volcano is one example of adapting these techniques to Earth science data. Figure 9 shows the completed graphic, including scales and annotations, as it appeared on the Earth Observatory.

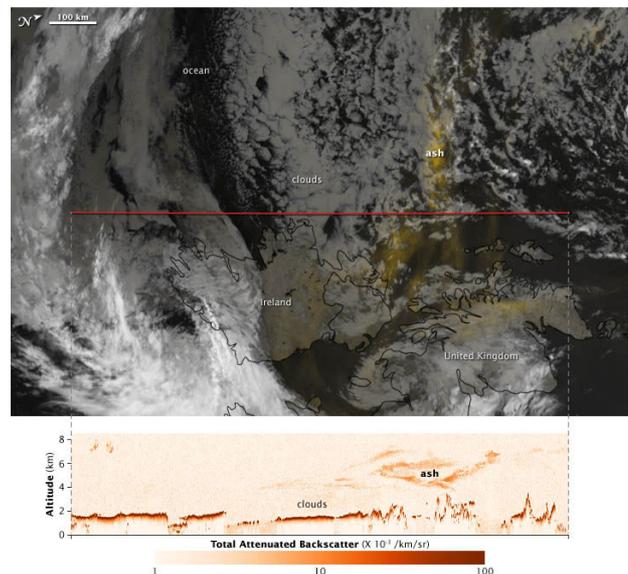


Figure 9: Combined CALIPSO/MODIS ash image.

References

Brewer, C., and Harrover, M., (n.d.) Color Brewer 2.0.
<http://colorbrewer2.org/>

Ellrod, G., (May 2010), National Weather Association,
"Remote Sensing of Volcanic Ash",
<http://www.nwas.org/committees/rs/volcano/ash.htm>

NASA Ames Research Center, Color Usage Research Lab.,
(n.d.) "Using Color in Information Display Graphics,"
<http://colorusage.arc.nasa.gov/>

Phipps, M. and Rowe, S., "Seeing satellite data," *Public Understanding of Science*, vol 19, 3: p.p. 311-321, May 2010.

Rogowitz, B. and Treinish, L., "Data Visualization: The End of the Rainbow," *IEEE Spectrum*, 35, 12: p.p. 52-59, December 1998.

Simmon, R. and Riebeek, H., 2010. NASA Earth Observatory.
"Nighttime Ash Tracking with CALIPSO,"
<http://earthobservatory.nasa.gov/IOTD/view.php?id=44052>

Tufte, E. 1997., *Visual Explanations*. Graphics Press, Cheshire, Connecticut.

Tufte, E., *The Visual Display of Quantitative Information: Second Edition*, Graphics Press, Cheshire, Connecticut. 2001.

Ware, C. 2004. *Information Visualization, Perception for Design: Second Edition*. Morgan Kaufmann.