POINTS OF VIEW

Into the third dimension

Three-dimensional visualizations are effective for spatial data but rarely for other data types.

When working with high-dimensional data, it may be tempting to choose a three-dimensional (3D) spatial visualization over a two-dimensional (2D) 'flat' representation because it allows us an additional data dimension. However, because quantitative, categorical and relational data are often not representing spatial relationships, plotting them in 3D space adds a level of visual complexity that often makes the data more difficult to understand. It therefore can be more effective to plot these data on a 2D plane and rely on nonspatial graphical encodings to represent additional dimensions.

For certain types of data, 3D spatial visualization is the best choice. For example, X-ray crystallography data describe the location of atoms in a molecule and thus characterize something that is inherently spatial. By visualizing the organization of these atoms in 3D space, we can reveal the molecular structure. Spatial data lend themselves to visual representations that reflect the 3D location information of the measurements—often crucial for the interpretation of the data (**Fig. 1**).

Two-dimensional projections of objects use visual depth cues to represent a third dimension. The strongest visual cue indicating depth is partial occlusion, in which one object hides parts of another. Another depth cue is the perspective created by converging parallel lines, which enables us to estimate the distances of objects from a certain vantage point. These depth cues are essential to depicting 3D objects on 2D displays (**Fig. 1**).

When data are plotted in 3D space, the visual cues needed to indicate depth may interfere with commonly used visual encodings. For example, the height or length of objects can be distorted by perspective, making it difficult to judge the scale of elements in a plot. Unavoidably, data objects in the foreground will interfere with the visibility of elements further from the viewer (**Fig. 2a**). When color is used to represent quantities, shading or shadows cast onto objects as depicted by the computer software can lead to further ambiguities.

The choice between a planar and a spatial representation should depend on whether the interference between visual encodings and depth cues constitutes an acceptable tradeoff given the goals of the visualization. Abstract data such as those generated for gene
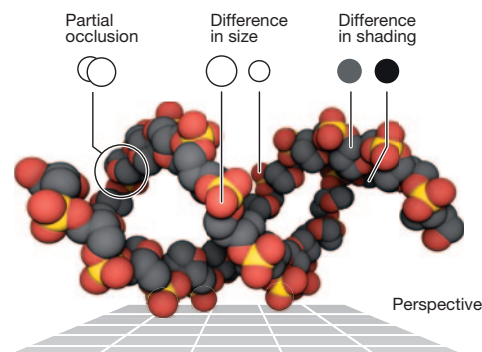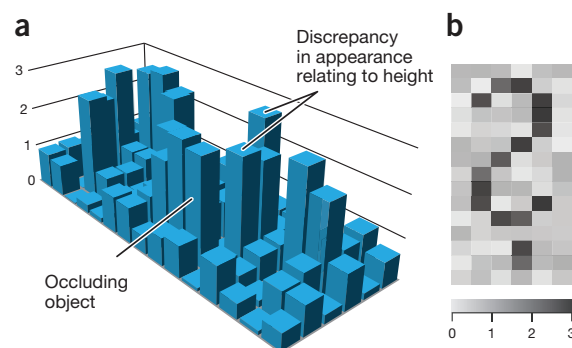

Figure 2 | Three-dimensional representation of abstract data. (**a**) Data occlusion and interference of visual encodings with depth cues can be problematic in three-dimensional space. (**b**) The same data as in **a** plotted as a two-dimensional heat map.

expression or biological networks do not generally benefit from 3D spatial representations and are most useful when plotted using techniques that do not require depth cues.

In most instances, high-dimensional data can be reliably and efficiently visualized with representations that place elements on a 2D plane and use size or color to encode further dimensions of the data (**Fig. 2b**). If one of the data dimensions is categorical and there are only a few categories, shapes can be used to encode the categories. Many general data visualization approaches are available to effectively represent multidimensional data on a plane. For example, a matrix of scatter plots each showing pairwise combinations of variables from a high-dimensional data set can productively reveal correlations. Similarly, heat maps and parallel coordinate plots[1] are useful techniques for plotting multidimensional data on a plane. If some information loss is acceptable, dimensionality reduction methods such as principal component analysis or multidimensional scaling can be used to obtain a 2D representation of a high-dimensional data set.

When a 3D spatial representation is chosen, the impact of occlusion should be minimized. In interactive visualizations, animated rotation of objects of interest is a common solution to show hidden surfaces. Additionally, semitransparent surfaces can be used to allow the viewer to look through or into objects, but this practice typically creates unintended visual artifacts, especially when color is also employed. When labels are required to describe 3D scenes, it is preferable to place them after the projection to the 2D display has been computed. If placed directly in the 3D scene, the labels may be distorted by the projection and become difficult or impossible to read.

Effective 3D spatial visualizations can be created by taking the properties of the data into account and applying depth cues that best support the visualization's communication goals. If such visualizations are applied to abstract data, the resulting visualization needs to offer significant benefits over nonspatial representations of the data.

Nils Gehlenborg & Bang Wong

1. Gehlenborg, N. & Wong, B. *Nat. Methods* **9**, 213 (2012).

Nils Gehlenborg is a research associate at Harvard Medical School and the Broad Institute. Bang Wong is the creative director of the Broad Institute and an adjunct assistant professor in the Department of Art as Applied to Medicine at The Johns Hopkins University School of Medicine.

Figure 1 | Space-filling model of the DNA backbone. Depth cues enable us to perceive two-dimensional images as three-dimensional objects.